



ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Επιχειρησιακή Ευφυΐα/Αναλυτική με Χρήση Δεδομένων ERP
και CRM Συστημάτων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της
Σκότη-Χριστοπούλου Φωτεινής-Δήμητρας

Επιβλέπων : Δρ. Λουκής Ευριπίδης, Καθηγητής

Μέλη εξεταστικής επιτροπής: Δρ. Ιωάννης Χαραλαμπίδης, Καθηγητής
Δρ. Νίκη Κυριακού, Μεταδιδακτορική Ερευνήτρια

Σάμος, Ιούλιος 2020

Πρόλογος και ευχαριστίες

Η παρούσα διπλωματική Εργασία εκπονήθηκε κατά την θερινή περίοδο του Ακαδημαϊκού Έτους 2020, στα πλαίσια του Προγράμματος Προπτυχιακών Σπουδών του Τμήματος Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων του Πανεπιστημίου Αιγαίου. Η εργασία πραγματοποιήθηκε υπό την επίβλεψη του Δρ. Λουκή Ευριπίδη, Καθηγητή του Τμήματος Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων, σε συνεργασία με τον κ. Κουταβίδη Σέργιο, Υπεύθυνο του τμήματος Business Intelligence and Data Analytics της Price Waterhouse Coopers Ελλάδος. Αντικείμενο της εργασίας αυτής είναι η Επιχειρηματική Ευφυΐα (Business Intelligence-BI), τα χαρακτηριστικά της καθώς και τα οφέλη που προσφέρει στο χώρο των επιχειρήσεων. Στα πλαίσια της εργασίας θα αναπτύξουμε ένα πλήρες σενάριο BI κάνοντας χρήση δεδομένων από ERP και CRM συστήματα.

Στο σημείο αυτό, θα ήθελα να ευχαριστήσω όλους όσους συνέβαλαν στην εκπόνηση της διπλωματικής μου εργασίας. Οφείλω να εκφράσω τις θερμές μου ευχαριστίες, προς τον επιβλέποντα της εργασίας, Καθηγητή Δρ. Λουκή Ευριπίδη, για την καθοδήγησή του, αλλά κυρίως για την ευκαιρία που μου έδωσε να ασχοληθώ και να εμβαθύνω τις γνώσεις μου σε ένα τόσο ενδιαφέρον και καινοτόμο θέμα Επίσης, τον κ. Κουταβίδη Σέργιο, για την εύρεση του θέματος καθώς και την καθοδήγησή και πολύτιμη βοήθεια του σε όλα τα στάδια αυτής της εργασίας καθώς και την PWC Ελλάδος για την παραχώρηση των δεδομένων που χρησιμοποιήθηκαν σε αυτή την διπλωματική εργασία. Εν συνεχεία, να ευχαριστήσω την εταιρεία μου NousPratIT, η οποία μου παραχώρησε όλους τους απαραίτητους πόρους για την ολοκλήρωση της εργασίας αυτής. Χωρίς τη συμπαράσταση και συνεχή βοήθεια όλων, η ολοκλήρωση αυτής της εργασίας δεν θα ήταν δυνατή. Τέλος, ευχαριστώ θερμά τους φίλους μου, για την κατανόηση και συμπαράσταση που έδειξαν ολόκληρη την περίοδο εκπόνησης της εργασίας αυτής.

© 2020

της

ΣΚΟΤΗ-ΧΡΙΣΤΟΠΟΥΛΟΥ ΦΩΤΕΙΝΗΣ-ΔΗΜΗΤΡΑΣ

Τμήμα Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων

ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

Πίνακας Περιεχομένων

1	Εισαγωγή	11
2	Επιχειρησιακή Αναλυτική	13
2.1	Διαχείριση Δεδομένων	13
2.2	Ποιότητα Δεδομένων	14
2.3	Μοντελοποίηση Δεδομένων	15
2.4	Διακυβέρνηση Δεδομένων	16
2.5	Αποθήκη Δεδομένων (Data Warehouse)	18
2.6	Επιχειρησιακή Ευφυΐα /Προγνωστική Αναλυτική.....	19
3	Cloud Computing	21
3.1	Εισαγωγή στο Cloud Computing	21
3.2	Πλεονεκτήματα	22
3.2.1	Cloud Data Warehouse.....	23
3.2.2	Business Intelligence Cloud.....	24
4	Μεθοδολογία.....	26
4.1	Τεχνολογία	26
4.2	Σχεδιαστικές Επιλογές (Εργαλεία και Τεχνολογίες).....	27
4.3	ERP, CRM εισαγωγή δεδομένων.....	29
4.4	Επεξεργασία Δεδομένων στο Cloud	31
4.5	Κατασκευή Τελικής Βάσης.....	48
4.6	Import Δεδομένων στο Power BI.....	51
4.7	Οπτικοποίηση και KIP's	54
4.7.1	Sales Analytics	55
4.7.2	CRM Analytics	59
4.8	Predictive Modeling and Advanced Analytics.....	64
5	Αποτελέσματα Ανάλυσης.....	71
5.1	Deployment στο σύννεφο	71
5.2	Διαδραστικότητα Αναφορών	74
5.3	Insights	75
6	Συμπεράσματα.....	79
	Βιβλιογραφία.....	81
	Παράρτημα I Κώδικας σε sql	83

Λίστα Εικόνων

Εικόνα 2.1: Σχήμα Αστέρα.....	16
Εικόνα 2.2: Διακυβέρνηση και Διαχείριση	17
Εικόνα 2.3: Data Warehouse	18
Εικόνα 2.4: Διαδικασία BI	20
Εικόνα 4.1: Cloud Data Warehouse & BI	26
Εικόνα 4.2: Microsoft Azure	28
Εικόνα 4.3: Sql Server.....	28
Εικόνα 4.4: Power BI	29
Εικόνα 4.5: Δεδομένα CRM.....	30
Εικόνα 4.6: Δεδομένα ERP	31
Εικόνα 4.7: Azure Sign Up.....	32
Εικόνα 4.8: Azure Data Factory	33
Εικόνα 4.9: Creation of Azure Data Factory	34
Εικόνα 4.10: Azure Storage.....	35
Εικόνα 4.11: Creation of Azure Storage.....	36
Εικόνα 4.12: Data Factory (Author & Monitor).....	36
Εικόνα 4.13: Data Factory Link Service	36
Εικόνα 4.14: New Linked Service.....	37
Εικόνα 4.15: New Data Set	38
Εικόνα 4.16: Δημιουργία Sql Server	39
Εικόνα 4.17: Import Database	39
Εικόνα 4.18: Table Mapping	40
Εικόνα 4.19: SSMS Connection to Sql Server	41
Εικόνα 4.20: SSMS Import File	41
Εικόνα 4.21: Choose Data Source	42
Εικόνα 4.22: Select Destination	42
Εικόνα 4.23: Select Source Tables and Views	43
Εικόνα 4.24: Import of CRM into Sql Server.....	44
Εικόνα 4.25: Linked Service CRM	45
Εικόνα 4.26: Κατασκευή Τελικής Βάσης.....	45
Εικόνα 4.27: Final Linked Services.....	46
Εικόνα 4.28: Pipeline for CRM	47
Εικόνα 4.29: Pipeline Completion.....	47
Εικόνα 4.30: Deployment Pipeline.....	47
Εικόνα 4.31: Τελική Δομή Βάσης	48
Εικόνα 4.32: Όψη Sales View	49
Εικόνα 4.33: Πίνακας CRM	50
Εικόνα 4.34: CRM Data	50

Εικόνα 4.35: CRM Data breakdown	51
Εικόνα 4.36: Όψη CRM Final	51
Εικόνα 4.37: Import Data to Power BI	52
Εικόνα 4.38: Connection to Server.....	52
Εικόνα 4.39: Transform/Load Data Power BI.....	53
Εικόνα 4.40: Sales Report	56
Εικόνα 4.41: Κατασκευή KPI.....	57
Εικόνα 4.42: Αλλαγή αλγορίθμου	57
Εικόνα 4.43: Αλλαγή μορφοποίησης	58
Εικόνα 4.44: Split Columns.....	58
Εικόνα 4.45: Διάγραμμα Revenue and Quantity	59
Εικόνα 4.46: CRM Analytics Report.....	60
Εικόνα 4.47: New measurement.....	61
Εικόνα 4.48: Dax Win Rate.....	61
Εικόνα 4.49: Dax Loss Rate	62
Εικόνα 4.50: Won Revenue.....	62
Εικόνα 4.51: Lost Revenue	62
Εικόνα 4.52: Month Measurement	62
Εικόνα 4.53: Filter Top 5	63
Εικόνα 4.54: Prediction and Analytics Report	65
Εικόνα 4.55: Probability	65
Εικόνα 4.56: Κατασκευή measurement για pipeline	66
Εικόνα 4.57: Trend Line Creation.....	67
Εικόνα 4.58: Trend Line.....	67
Εικόνα 4.59: Πωλήσεις 2013.....	68
Εικόνα 4.60: Forecasting.....	68
Εικόνα 4.61: Forecasting Upper-Lower	69
Εικόνα 4.62: Measurements για what if ανάλυση	69
Εικόνα 4.63: What if Analysis.....	70
Εικόνα 4.64: What if Analysis II.....	70
Εικόνα 5.1: Publish Report.....	71
Εικόνα 5.2: Επιλογή destination folder	72
Εικόνα 5.3: Publishing to Power BI	72
Εικόνα 5.4: Publishing to Power BI Success.....	72
Εικόνα 5.5: Power BI Dashboard	73
Εικόνα 5.6: Αρχικό Report CRM	74
Εικόνα 5.7: Επιλογή Salesman στο Crm Report	74
Εικόνα 5.8: Insight of online orders	75
Εικόνα 5.9: Correlation of online orders	76
Εικόνα 5.10: ClientId and Salesman	77
Εικόνα 5.11: Salesman and Status.....	78

Λίστα Πινάκων

Πίνακας 1: Βάση δεδομένων vs Αποθήκη Δεδομένων	19
--	----

Ακρωνύμια

BI	Business Intelligence
ETL	Extract-Load-Transform
PWC	Price Waterhouse Coopers
DSS	Decision Support Systems
SSMS	SQL Server Management Studio
DW	Data Warehouse
OTPL	Online Transactional processing
OLAP	Online Analytical processing
ERP	Enterprise resource planning
CRM	Customer relationship management
MIS	Management Information System
SaaS	Software as a Service
PaaS	Platform as a Service
IaaS	Infrastructure as a Service
DaaS	Desktop as a Service

Περίληψη

Η Επιχειρησιακή νοημοσύνη (EN) είναι ένας όρος της πληροφορικής ο οποίος εμπεριέχει όλα τα τμήματα και τις διεργασίες μιας επιχείρησης με σκοπό την ερμηνεία του τρόπου λειτουργίας της. Τα τελευταία χρόνια όλο και περισσότερες επιχειρήσεις στρέφονται στη χρήση εργαλείων επιχειρησιακής ευφυΐας προκειμένου να βελτιώσουν τις στρατηγικές τους αποφάσεις και να μπορέσουν να διαφοροποιηθούν και να εξελιχθούν. Ο κύριος στόχος της παρούσας εργασίας είναι να γίνει κατανοητή η συνεισφορά των EN συστημάτων με χρήση επιχειρησιακών δεδομένων από συστήματα προγραμματισμού πόρων και συστήματα διαχείρισης πελατειακών σχέσεων καθώς και των καινοτομιών που αυτά προσφέρουν όσον αφορά τον τρόπο διοίκησης των επιχειρήσεων. Η σωστή αξιοποίηση των δεδομένων των παραπάνω συστημάτων μπορεί να οδηγήσει στην εξαγωγή χρήσιμων πληροφοριών οι οποίες σε συνδυασμό με την χρήση EN συστημάτων μπορούν να δώσουν μια καλύτερη κατανόηση της επιχείρησης. Κάθε έργο επιχειρησιακής ευφυΐας ξεκινάει με την κατανόηση του τρόπου λειτουργίας της επιχείρησης καθώς και των στόχων που θέλει να επιτύχει. Εφόσον καθοριστούν οι στρατηγικοί στόχοι της επιχείρησης είναι εύκολο να προσδιοριστούν οι απαιτήσεις οι οποίες περιλαμβάνουν οικονομικούς δείκτες που μέλη της διοίκησης της επιχείρησης πιθανός να θέλουν να παρακολουθούν. Ακολουθεί η συλλογή των δεδομένων από τα διάφορα συστήματα της επιχείρησης και ο μετασχηματισμός αυτών σε μορφή κατάλληλη για την εφαρμογή αλγορίθμων ανάλυσης. Έπειτα γίνεται ανάλυση των δεδομένων δημιουργώντας αναφορές που περιέχουν συγκεντρωτικές πληροφορίες που η εκάστοτε επιχείρηση προτιμά να παρακολουθεί. Από αυτές τις αναφορές οι οποίες αποτελούν και το τελικό προϊόν ενός έργου EN, τα διοικητικά στελέχη αποκτούν εικόνα για την τρέχουσα πορεία της επιχείρησης εντοπίζοντας αδυναμίες που πρέπει να διορθώσουν καθώς και ευκαιρίες που μπορούν να αξιοποιήσουν στο μέλλον.

Λέξεις Κλειδιά: *EN, Συστήματα Προγραμματισμού Πόρων, Συστήματα Διαχείρισης Πελατειακών Σχέσεων*

Abstract

Business intelligence (BI) is an IT term that encompasses all parts of a business and the processes for interpreting its operation. In recent years, more and more companies are turning to the use of BI tools in order to improve their strategic decisions and help them move forward and evolve. The main goal of our research is to understand the contributions of BI systems using enterprise data from Resource Planning Systems (ERP) and Customer Relationship Management Systems (CRM) and the innovations they offer regarding the management of a company. The correct use of the enterprise data that ERP and CRM systems contain can lead to the collection of useful information that in combination with the use of BI systems can give a better understanding of the business. Every business intelligence project begins with understanding how the business operates and the goals that the business wants to achieve. Once the strategic goals of the business are clearly defined it is easy to identify the requirements, for example the economic indicators that the business management team may want to follow. Following up, we will present the collection of data from the various systems of the company and showcase their transformation into a format suitable for the application of analysis algorithms. Data are then being analyzed and we create reports that contain aggregate information that the company wants to monitor. From these interactive reports which are the final product of a BI project, give the executives a direct picture of the current course of the business, highlighting weaknesses that need to be corrected or even opportunities that they can take advantage of in the future.

Keywords: *BI, ERP, CRM*

1

Εισαγωγή

Ο όρος Business Intelligence εμφανίστηκε σαν ορολογία το 1989 για να περιγράψει έννοιες και μεθόδους που στόχο είχαν να βελτιώσουν τον τρόπο με τον οποίο οι επιχειρήσεις έπαιρναν αποφάσεις, συνεπώς στόχο είχε να υποστηρίξει την λειτουργία των συστημάτων υποστήριξης αποφάσεων (DSS). Η ανάγκη για έξυπνες επιχειρησιακές μεθόδους άρχισε να αυξάνεται τα τελευταία χρόνια, καθώς ο όγκος των δεδομένων που οι επιχειρήσεις διαχειρίζονται γίνεται ολοένα και μεγαλύτερος ενώ ταυτόχρονα αυξάνεται και η πολυπλοκότητα τους. Έτσι, δημιουργείται η ανάγκη για ευφυή πληροφοριακά συστήματα, τα οποία μπορούν να διαχειριστούν μεγάλο όγκο δεδομένων από πολλαπλές πηγές, να τα μετασχηματίσουν και να εξάγουν χρήσιμες πληροφορίες από αυτά. Οι πληροφορίες αυτές μπορούν είτε να προβλέψουν ρίσκα που ενδεχομένως να βλάψουν την πορεία της επιχείρησης είτε να αναγνωρίσουν ευκαιρίες στο ευρύτερο περιβάλλον της επιχείρησης. Με αυτό τον τρόπο λοιπόν η επιχείρηση μπορεί να ακολουθήσει καλύτερη στρατηγική με αποτέλεσμα να λειτουργεί καλύτερα, να αυξήσει τα κέρδη της και άρα να γίνει πιο ανταγωνιστική έναντι των αντιπάλων της. [1]

Στο 2ο κεφάλαιο θα γίνει αναφορά σε κάποιες βασικές έννοιες όπως είναι η επιχειρησιακή ευφυΐα και αναλυτική που αποτελούν και θέμα της διπλωματικής αυτής καθώς και σε κάποιες διαδικασίες οι οποίες είναι απαραίτητες για την εφαρμογή των παραπάνω. Τέτοιες διαδικασίες είναι η διαχείριση και διακυβέρνηση των δεδομένων που αποτελούν πολύ σημαντικό κεφάλαιο καθώς πρέπει να γίνει αντιληπτή η σημασία της σωστής επιλογής, συλλογής καθώς και διαχείρισης των δεδομένων που θα χρησιμοποιηθούν στα BI συστήματα. Θα αναλυθούν οι διάφορες τεχνικές μοντελοποίησης δεδομένων που υπάρχουν με έντονη έμφαση στην πιο χρησιμοποιούμενη μέθοδο μοντελοποίησης και αποθήκευσης που χρησιμοποιείται στα BI έργα, τα Data Warehouses.

Στο 3ο κεφάλαιο θα γίνει αναφορά στην cloud τεχνολογία, τα οφέλη χρήσης της και τον τρόπο με τον οποίο οι cloud τεχνολογίες ενισχύουν και υποστηρίζουν την χρήση της επιχειρησιακή ευφυΐας. και των Data warehouses.

Στο 4ο κεφάλαιο ξεκινάει η ανάλυση του πρακτικού μέρους της εργασίας. Θα περιγράφει ο τρόπος εξαγωγής και μοντελοποίησης των δεδομένων μέσω της ETL διαδικασίας (Extract-Transform-Load) με χρήση Cloud (Azure Data Factory), κατά την οποία ερχόμαστε σε επαφή με τα δεδομένα της επιχείρησης. Στη συνέχεια θα παρουσιαστεί ο τρόπος με τον οποίο κατασκευάστηκε τελική συγκεντρωτική βάση δεδομένων με χρήση του SQL Management Studio (SSMS), στην οποία θα εφαρμοστεί μετέπειτα η επιχειρησιακή αναλυτική. Μετά την κατασκευή του της βάσης αυτής θα εφαρμοστούν αλγόριθμοι αναλυτικής μέσω του εργαλείου οπτικοποίησης Microsoft Power BI, και θα κατασκευαστούν οι διαδραστικές αναφορές των οποίων τελικός χρήστης είναι τα διοικητικά στελέχη της επιχείρησης.

Στο 5ο κεφάλαιο θα δειχθεί ο τρόπος διαμοίρασης των αναφορών στο cloud καθώς και τα στοιχεία που μπορεί να εξάγει ένας χρήστης από τις αναφορές που κατασκευάστηκαν εκμεταλλευόμενος την διαδραστικότητα που αυτές προσφέρουν. Ακόμα θα δούμε κάποια insights που προσφέρει το Power BI, θα μιλήσουμε δηλαδή για την αυτόματη ανάλυση των δεδομένων που εφαρμόζει στα δεδομένα που έχουμε κάνει upload στην εφαρμογή και θα συγκρίνουμε τα αποτελέσματα με αυτά της ανάλυσής μας.

Τέλος στο 6ο κεφάλαιο θα εξάγουμε συμπεράσματα σχετικά με την έρευνα που υλοποιήθηκε. Θα αναφερθούν οι λόγοι για τους οποίους πρέπει κάθε επιχείρηση να αξιοποιεί τα δεδομένα της και να εξάγει πληροφορίες από αυτά καθώς και με ποιο τρόπο τα συστήματα επιχειρησιακής ευφυΐας χρησιμοποιώντας τα κατάλληλα δεδομένα μπορούν να δώσουν μια καλύτερη εικόνα για την επιχείρηση και τον τρόπο λειτουργίας της και να επεκτείνουν τις δυνατότητες των συστημάτων υποστήριξης αποφάσεων στον τρόπο διοίκησης και διαμόρφωσης στρατηγικής.

2

Επιχειρησιακή Αναλυτική

2.1 Διαχείριση Δεδομένων

Όπως αναφέρθηκε και προηγουμένως οι πληροφορίες που μπορεί να εξάγει μια επιχείρηση από τα δεδομένα, είναι αυτές που μπορούν να την οδηγήσουν στο να πάρει καλύτερες αποφάσεις και κατ' επέκταση να γίνει πιο κερδοφόρα. Ωστόσο, η διαχείριση ψηφιακών δεδομένων είναι μια διαδικασία που προϋποθέτει την ύπαρξη κάποιων εργασιών, πολιτικών διαδικασιών και πρακτικών και καλύπτει ένα μεγάλο εύρος [2]. Είναι πολύ σημαντικό να διαφυλάσσεται η ακεραιότητα της πληροφορίας τόσο κατά την συλλογή αυτής όσο και κατά την επεξεργασία και αποθήκευση της [2].

Στόχος της διαχείρισης δεδομένων είναι τόσο η άντληση των κατάλληλων πληροφοριών που θα αποδώσουν περισσότερες πληροφορίες, όσο και η κατάλληλη επεξεργασία. Η διαχείριση δεδομένων λοιπόν περιλαμβάνει από την διαδικασία εύρεσης των κατάλληλων δεδομένων για συλλογή έως και την διαχείριση της βάσης που θα περιέχει όλα τα υπό επεξεργασία δεδομένα. Είναι μια διαδικασία που απαιτεί γνώση τόσο του business όσο και τεχνική (διαχείριση βάσης δεδομένων).

Το πιο σημαντικό είναι να γνωρίζουμε τι δεδομένα διαθέτει στην κατοχή του κάθε οργανισμός και πως μπορεί να αξιοποιήσει καθένα από αυτά με σκοπό την επίτευξη στόχων. Τα δεδομένα πρέπει να ακολουθούν τις παρακάτω αρχές [2]:

- **Τα δεδομένα πρέπει να αντιμετωπίζονται σαν κεφάλαια της επιχείρησης.** Όπως κάθε άλλο κεφάλαιο της επιχείρησης έτσι πρέπει να αντιμετωπίζονται και τα δεδομένα.
- **Η κερδοφορία των δεδομένων πρέπει να εκφράζεται με οικονομικούς όρους.** Οι επιχειρήσεις πρέπει να εκτιμούν το κέρδος των δεδομένων τους και να δίνουν ιδιαίτερη σημασία στην διατήρηση της ποιότητας τους.

- **Η διαχείριση δεδομένων ισοδυναμεί με τη διατήρηση της ποιότητας.** Το σημαντικότερο για μια επιχείρηση είναι να διατηρεί την ποιότητα των δεδομένων της πάντα ανάλογα με τις απαιτήσεις που έχει η ίδια και οι συνεργάτες της.
- **Για την διαχείριση δεδομένων μεγάλο ρόλο παίζουν τα metadata.** Metadata ουσιαστικά είναι τα δεδομένα που σχετίζονται με τα δεδομένα που μας ενδιαφέρουν. (πελάτες, προϊόντα στην περίπτωση της πώλησης).

2.2 Ποιότητα Δεδομένων

Όπως αναφέρθηκε είναι πολύ σημαντικό να διαφυλαχτεί η ακεραιότητα και ποιότητα των δεδομένων που έχουμε σκοπό να χρησιμοποιηθούν για την ανάλυση της επιχείρησης διότι τα δεδομένα αυτά είναι που θα δώσουν μια εικόνα για την επιχείρηση την δεδομένη στιγμή. Για παράδειγμα αν μελετηθούν τα δεδομένα σχετικά με τις πωλήσεις ενός συγκεκριμένου προϊόντος μπορούν να εξαχθούν συμπεράσματα για την προτίμηση του κοινού σχετικά με το προϊόν και να οργανωθούν με βάση αυτά την στρατηγική της επιχείρησης μας.

Αν λοιπόν τα δεδομένα που θα χρησιμοποιηθούν στην ανάλυση δεν είναι έμπιστα η επιχείρηση μπορεί να καταλήξει ζημιωμένη. Μάλιστα σύμφωνα με υπολογισμούς της IBM [3], θέματα διαχείρισης της ποιότητας των δεδομένων, μόνο για τις ΗΠΑ, το έτος 2016 κόστισαν περί τα 3 δις.

Αν και είναι δύσκολο να εντοπιστούν με ακρίβεια τα κόστη που σχετίζονται με την χαμηλής ποιότητας δεδομένων, εκτιμάται ότι σχετίζονται με κόστη που προέρχονται από [3]:

- Διορθωτικές διαδικασίες
- Σύγκρουση συμφερόντων μέσα στην επιχείρηση
- Χαμηλή παραγωγικότητα
- Μη ικανοποίηση πελατών
- Έλλειψη δυνατότητας για πρωτοτυπία
- Πρόστιμα/διάφορα κόστη

Αυτό σημαίνει ότι υψηλής ποιότητας δεδομένα επιφέρουν τα αντίθετα αποτελέσματα μεταξύ αυτών υψηλή παραγωγικότητα, μειωμένα κόστη, αύξηση της παραγωγικότητας κλπ. Η διατήρηση της ποιότητας των δεδομένων όμως είναι μια διαδικασία που απαιτεί οργάνωση και αφοσίωση από την μεριά της εταιρείας.[3]

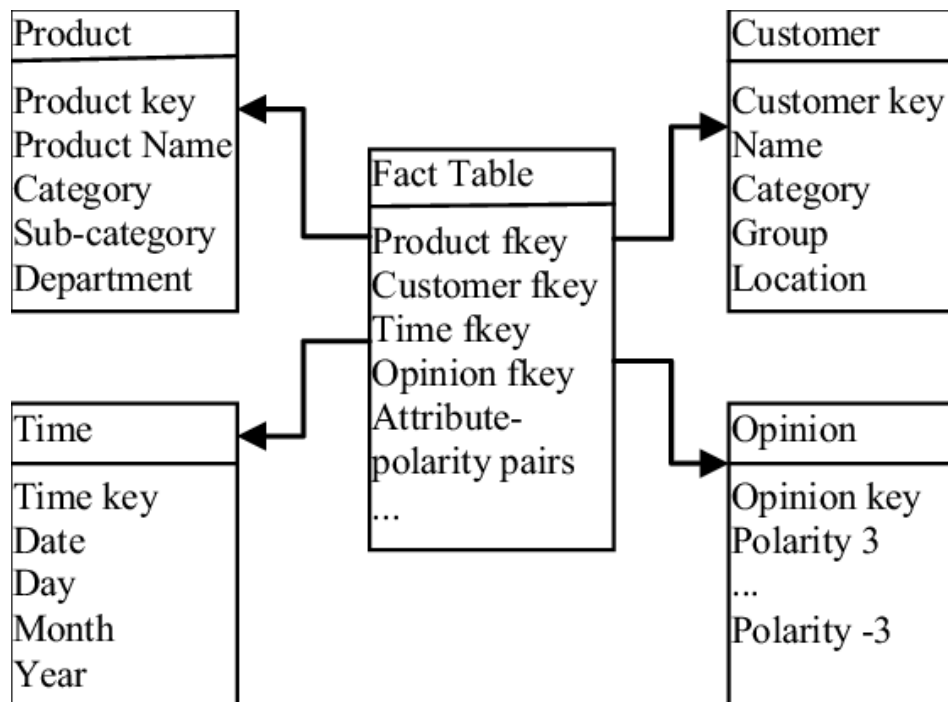
2.3 Μοντελοποίηση Δεδομένων

Η μοντελοποίηση αποτελεί και αυτή με την σειρά της ένα μέρος της διαδικασίας διαχείρισης δεδομένων και είναι η διαδικασία εύρεσης, ανάλυσης και καθορισμού των προδιαγραφών που πρέπει να ακολουθούν τα δεδομένα. Περιλαμβάνει την διαδικασία μορφοποίησης των δεδομένων σύμφωνα με ένα πρότυπο/μοντέλο το οποίο πρέπει να ακολουθεί τις προδιαγραφές της κάθε επιχείρησης. Όσο σημαντικό είναι λοιπόν να διατηρηθεί η ποιότητα των δεδομένων, άλλο τόσο είναι και η μετατροπή τους σε κατάλληλη μορφή ώστε να εξαχθούν εύκολα πληροφορίες και συμπεράσματα από αυτά. Υπάρχουν πολλοί τρόποι να μορφοποιηθούν τα δεδομένα, κάποιοι από τους πιο γνωστούς είναι η σχεσιακή, αυτή των διαστάσεων, η αντικειμενοστραφής, των γεγονότων (fact-table) και η No SQL. [3] Τέτοιου είδους μοντέλα διαχωρίζουν την πληροφορία σε τρία επίπεδα πληροφορίας, την θεωρητική, λογική και φυσική. Καθένα από αυτά τα μοντέλα περιλαμβάνει οντότητες, σχέσεις, γεγονότα, κλειδιά και γνωρίσματα. Στόχος είναι η κατάλληλη επιλογή μοντέλου ανάλογα με τις ανάγκες της επιχείρησης, καθώς και η συντήρηση του σε βάθος χρόνου.

Το μοντέλο δεδομένων με το οποίο θα αναλυθεί στην επιχειρησιακή ευφυΐα είναι το μοντέλο γεγονότων (fact-table) ή αλλιώς σχήμα αστέρα. Το σχήμα αστέρα είναι μια μορφή μοντελοποίησης που χρησιμοποιείται κατά κανόνα από αποθήκες δεδομένων (data warehouses). Έχει σαν προαπαιτούμενο την ταξινόμηση της πληροφορίας ως προς κάποια διάσταση ή γεγονός. Η δομή αστέρα αποτελείται από δυο μέρη, δηλαδή τους πίνακες διαστάσεων και τον πίνακα γεγονότων.

Οι πίνακες διαστάσεων είναι συνήθως πίνακες με μικρό αριθμό εγγραφών και περιέχουν την αναπαράσταση επιχειρηματικών οντοτήτων, δεδομένων δηλαδή που υπάρχει η δυνατότητα να μορφοποιηθούν. Τέτοια μπορεί να είναι προϊόντα, άτομα, μέρη και έννοιες, συμπεριλαμβανομένου και του ίδιου του χρόνου. Ο πιο διαδεδομένος πίνακας σε ένα σχήμα αστέρα είναι αυτός της διάστασης του χρόνου. [4]

Οι πίνακες γεγονότων χρησιμοποιούνται για την αποθήκευση μετρήσεων ή μέτρησης συμβάντων δηλαδή αποτελέσματα υπολογισμών και μπορούν να είναι παραγγελίες πωλήσεων, υπόλοιπα αποθεμάτων, συναλλαγματικές ισοτιμίες κ.λπ. Ένας πίνακας γεγονότων αποτελείται από στήλες που περιέχουν αριθμητικές τιμές και ξένα κλειδιά (foreign keys) που αναφέρονται σε πρωτεύοντα κλειδιά πίνακα διαστάσεων. Ένα τέτοιο παράδειγμα παρουσιάζεται στην παρακάτω εικόνα που αποτελεί σχήμα μιας βάσης αστέρα ενός καταστήματος λιανικής πώλησης.

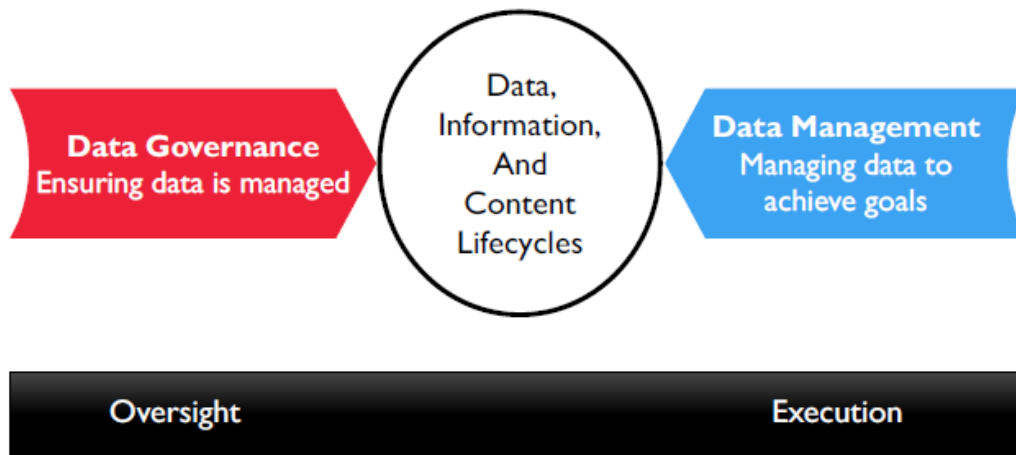


Εικόνα 2.1: Σχήμα Αστέρα

Από το παραπάνω σχήμα βάσης εξάγεται πολύ εύκολα η πληροφορία για το πιο προϊόν αγοράστηκε, από ποιόν πελάτη, ποια χρονική στιγμή και ποια ήταν η τελική εμπειρία από την αγορά. Η ίδια πληροφορία θα ήταν πολύ πιο χρονοβόρο να εξαχθεί από μια απλή σχεσιακή βάση δεδομένων όπου θα έπρεπε να κάνουμε πολλαπλά join μεταξύ πινάκων για να φέρουμε την ίδια πληροφορία. Όπως γίνεται αντιληπτό αυτό δεν θα ήταν πρόβλημα για μια μικρής έκτασης βάση αλλά όταν μιλάμε για μια μεγάλη επιχείρηση με Gigabytes ή ακόμα και Terabytes από Data ένα τέτοιο ερώτημα(query) θα είχε μεγάλο χρόνο εκτέλεσης και αυτό θα είχε αντίκτυπο στην απόκριση του συστήματός. [5]

2.4 Διακυβέρνηση Δεδομένων

Η διακυβέρνηση δεδομένων (Data Governance) περιλαμβάνει διαδικασίες που σχετίζονται με αποφάσεις που πρέπει να παρθούν σχετικά με την διαχείριση των δεδομένων μιας εταιρείας. Στόχο έχει να ελέγχει ότι τα δεδομένα διαχειρίζονται ακολουθώντας πάντα τις πρακτικές και τους στόχους της εταιρείας. Ενώ η διαδικασία διαχείρισης δεδομένων ασχολείται με το αν τα δεδομένα διαχειρίζονται με τέτοιο τρόπο ώστε να αποφέρουν κέρδος στην επιχείρηση η διακυβέρνηση δεδομένων ασχολείται με το τι ενέργειες πρέπει να ακολουθηθούν όσον αφορά στα δεδομένα και πώς τα μέλη της επιχείρησης πρέπει να τα διαχειρίζονται. [3]



Εικόνα 2.2: Διακυβέρνηση και Διαχείριση

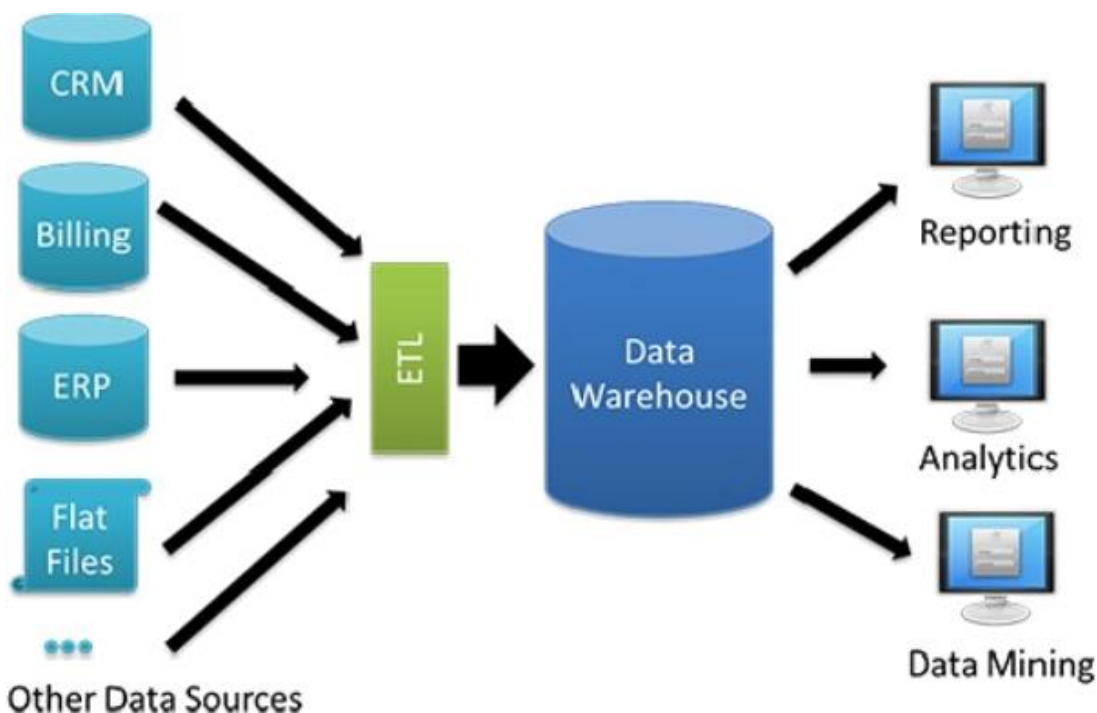
Η διακυβέρνηση δεδομένων περιλαμβάνει [3]:

- **Στρατηγική:** Καθορισμός στρατηγικής ως αναφορά τα δεδομένα και την διακυβέρνηση.
- **Πολιτική:** Διαμόρφωση πολιτικής ως αναφορά την διαχείριση των δεδομένων και των metadata τους, την χρήση τους, την ασφάλεια τους την ποιότητα καθώς και την ασφάλεια τους.
- **Πρότυπα και ποιότητα:** Ρύθμιση και επιβολή χρήσης προτύπων ως αναφορά την ποιότητα των δεδομένων και την εφαρμογή κατάλληλων αρχιτεκτονικών σε αυτά.
- **Συμμόρφωση:** Εξασφάλιση στο ό,τι η επιχείρηση θα συμμορφώνεται με τις απαιτήσεις κανονικότητας των δεδομένων.
- **Διαχείριση ζητημάτων:** Εντοπισμός και λύση οποιουδήποτε προβλήματος σχετικά με την ποιότητα των δεδομένων, την ασφάλεια, τα πρότυπα ή ακόμα και την ιδιοκτησία τους.
- **Αποτίμηση των δεδομένων:** Εδραίωση προτύπων και διαδικασιών με στόχο τον καθορισμό του κέρδους που μπορούν να επιφέρουν τα δεδομένα.
- **Έργα διαχείρισης δεδομένων:** Προσπάθεια βελτίωσης του τρόπου εφαρμογής της διαχείρισης δεδομένων.

2.5 Αποθήκη Δεδομένων (Data Warehouse)

Η έννοια του Data Warehouse (DW) εμφανίστηκε στις αρχές του 1980 ως μια τεχνολογία που θα ωθούσε τις εταιρείες να ενσωματώσουν τα δεδομένα τους από διάφορες πηγές σε ένα ενιαίο μοντέλο. Τα ενσωματωμένα αυτά δεδομένα υπόσχονταν να προσφέρουν διορατικότητα στις διαδικασίες της εταιρείας και να δημιουργήσουν νέες δυνατότητες αξιοποίησης δεδομένων με στόχο την καλύτερη λήψη αποφάσεων. [6] Τα πιο διαδεδομένα εργαλεία για την δημιουργία μιας αποθήκης δεδομένων είναι της IBM Info sphere, Oracle, Amazon Redshift κα. [6]

Η κατασκευή τους άρχισε το 1990 και άνοιξε με την ταυτόχρονη άνοδο του Business Intelligence. Πλέον στις μέρες μας είναι αναπόσπαστο κομμάτι των επιχειρήσεων όσον αφορά στον τρόπο αποθήκευσης και διαχείρισης των δεδομένων. Ένα DW είναι μια βάση δεδομένων non OTPL (Online Transactional processing), και είναι γνωστή ως OLAP (Online Analytical processing), η οποία λειτουργεί σαν ένα στρώμα πάνω από μια transactional βάση δεδομένων. Συνήθως χρησιμοποιεί το σχήμα αστέρα που έχουμε περιγράψει και διατηρεί ιστορικά δεδομένα έτσι ώστε να μπορεί να γίνει πιο εύκολη η χρήση predictive μοντέλων και analysis από κάποιο BI application.



Εικόνα 2.3: Data Warehouse

Τα πλεονεκτήματα ενός data warehouse σε σχέση με μια απλή βάση δεδομένων παρουσιάζεται αναλυτικά και στον παρακάτω πίνακα. [7]

	Βάση Δεδομένων	Data Warehouse
Ορισμός	Κάθε είδους δεδομένο, με στόχο την αποθήκευση, επεξεργασία ή ανάκτηση	Επίπεδο ενδιάμεσα σε βάσεις δεδομένων και εργαλείων ανάλυσης
Τύποι Δεδομένων	Excel, Flat Files, OLTP, XML	OLAP
Οργάνωση Δεδομένων	Πολύπλοκοι πίνακες, συνδέσεις μεταξύ πινάκων	Απλοποιημένη μορφή κατάλληλη για ανάλυση με μικρό χρόνο απόκρισης
Βελτιστοποίηση	Κατασκευασμένη για διάβασμα-γράψιμο δεδομένων, όχι βέλτιστη για ανάλυση	Ειδικά φτιαγμένη δομή για χρήση ανάλυσης και ανάκτηση πολλών δεδομένων
Αναφορές/Ανάλυση	Μπορεί να βγει οποιαδήποτε αναφορά με δεδομένα από την βάση αλλά απαιτεί γνώση	Χωρίς ιδιαίτερη γνώση κάποιος μπορεί να εφαρμόσει διαφόρους τύπους ανάλυσης

Πίνακας 1: Βάση δεδομένων vs Αποθήκη Δεδομένων

2.6 Επιχειρησιακή Ευφυΐα /Προγνωστική Αναλυτική

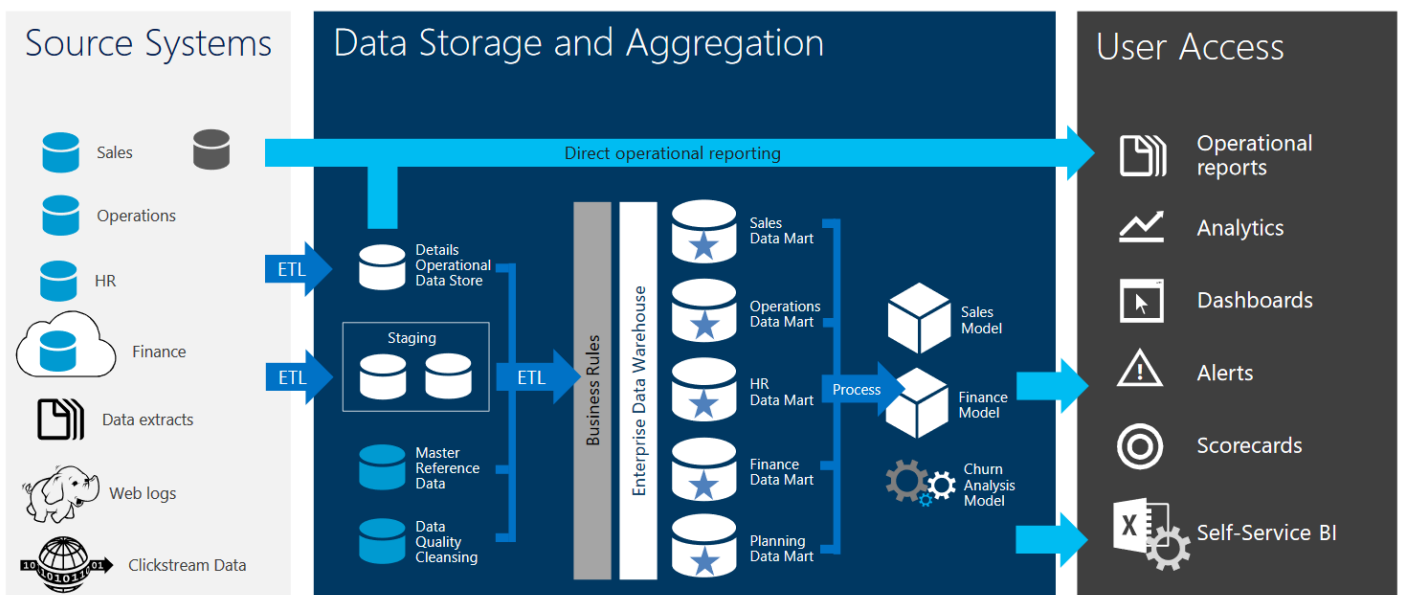
Τα τελευταία χρόνια σε όλους τους τομείς της οικονομίας παρατηρείται έντονος ανταγωνισμός. Αυτό αναγκάζει τις επιχειρήσεις να αναδιοργανώνονται συνεχώς και να προσαρμόζονται σε όποιες αλλαγές εμφανίζονται στο ευρύτερο περιβάλλον τους. Επομένως, κάθε επιχείρηση για να μπορέσει να είναι ανταγωνιστική και πιο κερδοφόρα πρέπει να είναι σε θέση να μπορεί να αναγνωρίσει απειλές καθώς και ευκαιρίες στο περιβάλλον της. Για αυτό τον λόγο οι επιχειρήσεις αναζητούν πρωτοποριακές λύσεις οι οποίες θα τις βοηθήσουν να αποκτήσουν ανταγωνιστικό πλεονέκτημα καθώς και να ανταπεξέλθουν σε οποιουσδήποτε επιχειρηματικούς κινδύνους μπορεί να προκύψουν [8].

Κάθε επιχείρηση χρησιμοποιεί πληθώρα πληροφοριακών συστημάτων και εργαλείων ώστε να μπορεί να είναι σε θέση να λειτουργήσει. Τέτοια πληροφοριακά συστήματα είναι αυτά της διαχείρισης επιχειρησιακών πόρων (ERP), της διαχείριση σχέσεων με τους πελάτες (CRM), συστήματα διοίκησης (MIS) καθώς και αυτά της Επιχειρησιακής Ευφυΐας (BI). Όπως καταλαβαίνουμε όλα αυτά τα συστήματα έχουν από πίσω τους ένα τεράστιο όγκο δεδομένων τον οποίο οι επιχειρήσεις πρέπει να είναι σε θέση να μπορούν να διαχειριστούν. Την ανάγκη αυτή για

διαχείριση μεγάλου όγκου δεδομένων από διαφορετικές πηγές έρχεται να λύσει η επιχειρησιακή Ευφυΐα.

Όπως έχει αναφερθεί το BI ξεκινάει με την ETL διαδικασία των δεδομένων της επιχείρησης, συνεχίζει με την δημιουργία ενός Data Warehouse και ολοκληρώνεται με την χρήση ενός BI Tool το οποίο δίνει την δυνατότητα να κάνουμε στατιστική ανάλυση των δεδομένων μας και να παράγουμε διαδραστικά διαγράμματα για τον τελικό χρήστη. Η ανάλυση αυτή μπορεί να περιλαμβάνει την εμφάνιση απλών στατιστικών δεικτών (Key Performance Indicators – KPIs) οι οποίοι δίνουν μια σαφή εικόνα για την επιχείρηση ή ακόμα και την πρόγνωση μελλοντικών γεγονότων και ανάλυση τάσεων μέσω προγνωστικής αναλυτικής. Η προγνωστική αναλυτική είναι ένα είδος προχωρημένης ανάλυσης με σκοπό την πρόβλεψη μελλοντικών γεγονότων κάνοντας χρήση ιστορικών δεδομένων και εφαρμόζοντας στατιστικών αλγορίθμων και ερωτημάτων σε αυτά.

Υπάρχουν διάφορα εργαλεία BI στην αγορά κάποια από τα πιο γνωστά είναι το Tableau, QlickView, Microsoft Power BI, SAS, SAP [9]. Τα εργαλεία αυτά παρέχουν μέσω διαδραστικών διαγραμμάτων μια πιο καθαρή εικόνα της επιχείρησης από αυτή που θα παίρναμε από ένα στατικό report ενός ERP συστήματος.



Εικόνα 2.4: Διαδικασία BI

3

Cloud Computing

3.1 Εισαγωγή στο Cloud Computing

Το υπολογιστικό νέφος (Cloud Computing) έκανε την εμφάνισή του για πρώτη φορά στις αρχές του 1950 σε εκπαιδευτικά ιδρύματα, αλλά και σε πολλές επιχειρήσεις. Με το πέρασμα των ετών και ιδίως μετά το 1970 έγινε ευρέως γνωστό [10]. Οι χρήστες του πρώτου νέφους αποκτούσαν πρόσβαση σε υπολογιστές μεγάλων υπολογιστικών και αποθηκευτικών δυνατοτήτων μέσω τερματικών εγκατεστημένων σε υπολογιστές με πολύ μικρή υπολογιστική ισχύ και χωρίς κάποια δυνατότητα αποθήκευσης [10]. Στις μέρες μας, υπολογιστικό νέφος είναι η διάθεση διάφορων πόρων ή υπηρεσιών απομακρυσμένα από τον χρήστη με πρόσβαση μέσω του διαδικτύου. Σε αυτές μπορεί να συγκαταλέγονται η παροχή υπολογιστικών πόρων, χώρων αποθήκευσης, δικτύωσης, εφαρμογών ή ακόμη και αναλυτική. Συνεπώς, παρέχονται στον χρήστη υπηρεσίες που του προσφέρουν διευκολύνσεις και ευελιξία- προσαρμοστικότητα σε καθημερινές εργασίες και καλύπτουν μεγάλο εύρος τομέων. Αυτές οι υπηρεσίες παρέχονται από εταιρείες που ονομάζονται πάροχοι Cloud και η κοστολόγηση γίνεται με βάση την χρήση.

Τα είδη υπηρεσιών του cloud computing, είναι τέσσερα [10]:

- **Software-as-a-Service (SaaS):** Υπηρεσία που παρέχεται μέσω internet, αντί για λογισμικό που θα επιβάρυνε τον πελάτη λόγω επιδιορθώσεων, εφαρμογών κτλ.
- **Platform-as-a-Service (PaaS):** Ο πελάτης είναι σε θέση να χρησιμοποιεί πλατφόρμες και λοιπά εργαλεία (πχ Java) αντί να αγοράσει άδειες λογισμικού για τις πλατφόρμες που θα του κόστιζαν περισσότερο.
- **Infrastructure-as-a-Service (IaaS):** Είναι απλές και βασικές υλικές συσκευές (εικονικά pc, διακοσμητές κτλ.) που είναι σε κάποιο φυσικά κεντρικό σημείο. Δίνεται η δυνατότητα

να χρησιμοποιούνται από το internet, κάνοντας χρήση συστημάτων ελέγχου πρόσβασης από οποιοδήποτε τερματικό ή συσκευή.

- **Desktop-as-a-Service (DaaS):** Αυτή η υπηρεσία παρέχει την υποδομή μίας εικονικής επιφάνειας εργασίας, η οποία φιλοξενείται από κάποιο πάροχο λογισμικού νέφους χρησιμοποιείτο μοντέλο μηνιαίας συνδρομής. Κάνει χρήση μιας αρχιτεκτονικής πολλαπλών μισθώσεων, άρα πρόκειται για μία εφαρμογή που εξυπηρετεί πολλούς χρήστες, τους λεγόμενους «ενοικιαστές». Ο πάροχος λύσεων λογισμικού σύννεφου έχει την ευθύνη να διαχειριστεί το cloud και την υποκείμενη υποδομή και φυσικά ο βαθμός της εξυπηρέτησης διαφέρει αναλογικά με τις ανάγκες που έχουν οι χρήστες. Ως τελικό αποτέλεσμα οι χρήστες έχουν πρόσβαση σε δεδομένα και εφαρμογές σχεδόν από κάθε συσκευή και οποιοδήποτε μέρος.

3.2 Πλεονεκτήματα

Τα κυριότερα πλεονεκτήματα είναι τα εξής [11]:

- **Κόστος απόδοσης:** Το υπολογιστικό σύννεφο είναι η αποδοτικότερη μέθοδος για τη χρήση, διατήρηση και αναβάθμιση των πληροφοριών του χρήστη. Η αγορά άδειας ενός παραδοσιακού on-premise λογισμικού σε πολλές περιπτώσεις κοστίζει πολύ περισσότερο. Η προσθήκη αδειών χρήσης για αρκετούς χρήστες μπορεί να αποδειχθεί πολύ ακριβή για αυτή την εγκατάσταση. Το σύννεφο από την άλλη πλευρά, διατίθεται σε πολύ φθηνότερη τιμή και λόγω αυτού μπορεί να μειώσει σημαντικά τα έξοδα πληροφορικής. Εκτός αυτού, υπάρχουν διάφορες επιλογές πληρωμής, πχ. Το να πληρώνεις όσο χρησιμοποιείς τις υπηρεσίες και να μην έχεις κάποιο πλάνο αγοράς και άλλες κλιμακωτές επιλογές στη διάθεσή μας.
- **Σχεδόν απεριόριστη αποθήκευση:** Η αποθήκευση πληροφοριών στο υπολογιστικό νέφος προσφέρει σχεδόν απεριόριστη αποθήκευση / χωρητικότητα.
- **Δημιουργία αντιγράφων ασφαλείας και αποκατάσταση:** Όταν τα δεδομένα αποθηκεύονται στο υπολογιστικό νέφος οι πάροχοι cloud κρατάνε και ένα backup ασφαλείας. Αυτό καθιστά την αποθήκευση σε ένα υπολογιστικό σύννεφο ασφαλέστερη από την αποθήκευση σε ένα φυσικό μέσο το οποίο μπορεί να υποστεί φυσική καταστροφή ανά πάσα στιγμή. Επιπροσθέτως, οι περισσότεροι πάροχοι υπηρεσιών cloud συνήθως δίνουν την δυνατότητα ανάκτησης πληροφοριών. Ως εκ τούτου, αυτό καθιστά το σύνολο της διαδικασίας backup και αποκατάστασης πολύ απλούστερη συγκριτικά με άλλες παραδοσιακές μεθόδους αποθήκευσης δεδομένων.

- **Αυτόματη ενσωμάτωση λογισμικού:** Στο σύννεφο, η ενσωμάτωση λογισμικού γίνεται αυτόματα. Αυτό σημαίνει ότι οι χρήστες δεν είναι απαραίτητο να καταβάλουν επιπλέον προσπάθειες ώστε να κάνουν προσαρμογές και να ενσωματώσουν τις εφαρμογές τους σύμφωνα με τις προτιμήσεις τους.
- **Εύκολη πρόσβαση στις πληροφορίες:** Αφού οι χρήστες εγγραφούν στο Υπολογιστικό Νέφος, δύνανται να έχουν πρόσβαση στις πληροφορίες από οπουδήποτε, αρκεί να υπάρχει σύνδεση στο διαδίκτυο. Αυτή η εύχρηστη λειτουργία επιτρέπει στους χρήστες να διαχειρίζονται τα δεδομένα τους ανεξαρτήτως ζώνης ώρας και γεωγραφικής θέσης.
- **Γρήγορη ανάπτυξη:** Μόλις ο χρήστης αποφασίσει να χρησιμοποιήσει το υπολογιστικό νέφος, το σύστημα είναι πλήρως λειτουργικό εντός μερικών λεπτών.
- **Παροχή νέων υπηρεσιών:** Συνεχώς προστίθενται νέες δυνατότητες και υπηρεσίες στα υπολογιστικά σύννεφα καλύπτοντας έτσι τις ανάγκες των χρηστών.

3.2.1 Cloud Data Warehouse

Το cloud Data warehouse είναι ένα βασικό στοιχείο οποιασδήποτε υποδομής ανάλυσης δεδομένων εντός της οποίας ένας πάροχος cloud εκχωρεί τον απαιτούμενο χώρο για την αποθήκευση των δεδομένων και την υπολογιστική ισχύ που απαιτείται για την επεξεργασία τους. Παρέχεται ως μια υπηρεσία που στόχο έχει την βελτίωση του τρόπου χρήσης της αναλυτικής και την ευκολία στη χρήση του τελικού χρήστη. Η ανάπτυξη του cloud είναι επί του παρόντος η υψηλότερη προτεραιότητα για τις εταιρείες (πάνω από το 50% των εταιρειών θεωρούν την χρήση του ως κρίσιμη ή πολύ σημαντική), γι' αυτό και τα τελευταία χρόνια, ολοένα και περισσότερες εταιρείες προχωράνε στη μεταφόρτωση των δεδομένων τους από τοπικούς διακομιστές στο cloud.

Οι αποθήκες δεδομένων που βασίζονται σε cloud προσφέρουν ελευθερία στις εταιρείες και δυνατότητα να επικεντρωθούν στη λειτουργία της επιχείρησής τους. Ακόμα, προσφέρουν μεγαλύτερες ταχύτητες κατά την διαδικασία ανάκτησης ή και μεταφόρτωσης των δεδομένων και ευκολία στην επεκτασιμότητα και διαχείριση αυτών. Μεταξύ άλλων προσφέρει [12]:

- **Πρόσβαση δεδομένων:** Η τοποθέτηση των δεδομένων τους στο cloud επιτρέπει στις εταιρείες να παρέχουν στους αναλυτές τους πρόσβαση σε δεδομένα σε πραγματικό χρόνο από πολλές πηγές, επιτρέποντάς τους να εκτελούν γρήγορα καλύτερα ανάλυση αυτών.
- **Επεκτασιμότητα:** Είναι πολύ πιο γρήγορο και λιγότερο ακριβό να κλιμακωθεί μια αποθήκη δεδομένων cloud από ένα σύστημα εγκατάστασης, επειδή δεν απαιτεί την αγορά νέου υλικού (και ενδεχομένως υπερβολικής ή ανεπαρκούς παροχής)

- Απόδοση: Μια αποθήκη δεδομένων cloud επιτρέπει την εκτέλεση ερωτημάτων πολύ πιο γρήγορα από ό, τι σε σχέση με μια παραδοσιακή αποθήκη δεδομένων εσωτερικής εγκατάστασης, με χαμηλότερο κόστος.

Κάθε ένας από τους μεγάλους δημόσιους προμηθευτές cloud προσφέρει τη δική του υπηρεσία αποθήκης δεδομένων στο εκάστοτε cloud του προμηθευτή: η Google προσφέρει το BigQuery, η Amazon το Redshift και η Microsoft διαθέτει το Azure SQL Data Warehouse. Για καθεμία από αυτές τις υπηρεσίες, ο προμηθευτής cloud ή ο πάροχος αποθήκης δεδομένων παρέχει τις ακόλουθες δυνατότητες "εκτός λειτουργίας" [12]:

- Αποθήκευση και διαχείριση δεδομένων: τα δεδομένα αποθηκεύονται σε ένα σύστημα αρχείων που βασίζεται σε σύννεφο.
- Αυτόματες αναβαθμίσεις: δεν υπάρχει έννοια "έκδοσης" ή αναβάθμισης λογισμικού.
- Διαχείριση χωρητικότητας: είναι πολύ εύκολο είτε να επεκτείνουμε τα δεδομένα μας είτε να τα μειώσουμε ανά πάσα στιγμή.

3.2.2 *Business Intelligence Cloud*

Οι εφαρμογές BI, οι οποίες περιλαμβάνονται στο Cloud κάνουν τα επιχειρηματικά δεδομένα να είναι πιο προσηνή από κάθε άλλη περίοδο. Οι εφαρμογές Cloud Business Intelligence (BI) φιλοξενούνται εντός ενός εικονικού δικτύου [13]. Χρησιμοποιούνται με στόχο την παροχή πρόσβασης σε δεδομένα σχετιζόμενα με το BI, όπως λόγου χάρη dashboards, δείκτες KPI κ.ά. σε διάφορους οργανισμούς. Γενικότερα οι επιχειρήσεις στρέφουν το ενδιαφέρον τους όλο και πιο πολύ σε εργαλεία που έχουν τη βάση τους στο cloud, όπως παραδείγματος χάριν οι εφαρμογές για τη διαχείριση σχέσεων μεταξύ πελατών (CRM), online διαμοίραση ή αποθήκευση αρχείων (Dropbox, Box) και λογισμικό υποστήριξης – helpdesk (UserVoice, Zendesk). Αυτή η τάση προσφέρει εργαλεία επιχειρηματικής ευφυΐας που περιλαμβάνουν την προσαρμοστικότητα και την προσβασιμότητα του cloud.

Το cloud computing συνδυάζεται ιδανικά με την επιχειρησιακή ευφυΐα. Το κομμάτι της επιχειρησιακής ευφυΐας σχετίζεται με την παροχή αξιόπιστων και ορθών πληροφοριών στους κατάλληλους ανθρώπους την σωστή στιγμή. Από την άλλη πλευρά το cloud computing προσφέρει έναν εύκολο, αλλά ταυτόχρονα και ευέλικτο τρόπο πρόσβασης στις εφαρμογές BI. Οι εφαρμογές του Cloud BI διακατέχονται από πολλά πλεονεκτήματα, όπως το ότι είναι προσβάσιμες σε

πληθώρα συσκευών και προγράμματα περιήγησης στο διαδίκτυο. Έτσι, λοιπόν, ξεπερνιούνται όλα τα εμπόδια λογισμικού.

Όπως προαναφέρθηκε το Cloud BI προσφέρει βασικά πλεονεκτήματα συγκριτικά με άλλες εφαρμογές οι οποίες εκτελούνται on-site. Κάποια από αυτά είναι τα εξής [13]:

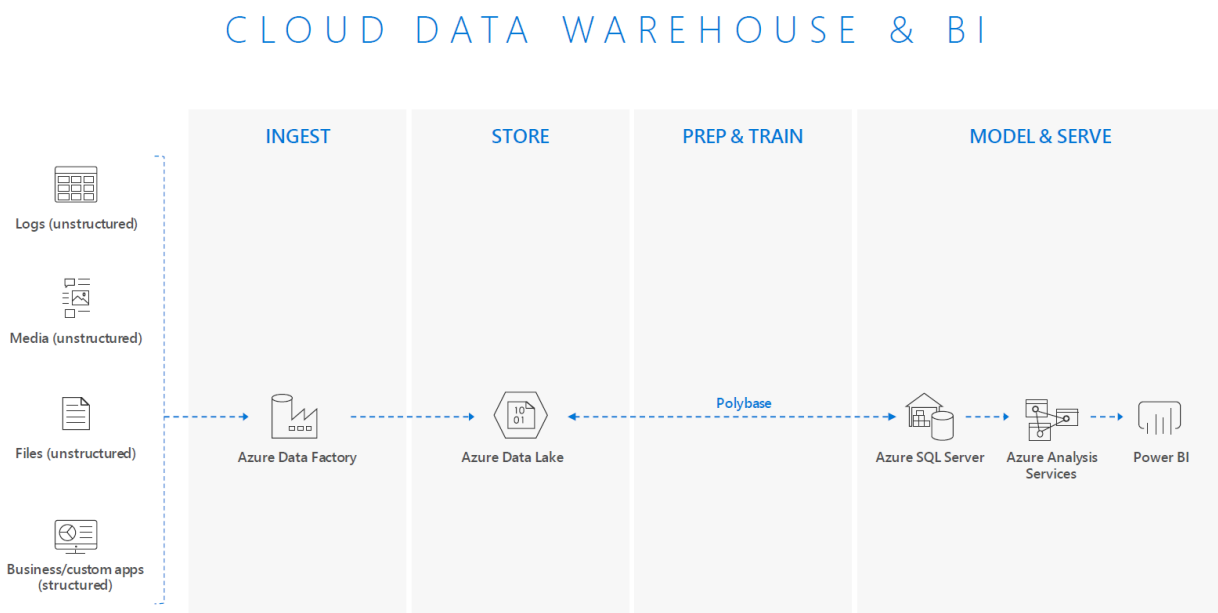
- Ευκολία στον τρόπο χρήσης
- Οι Cloud BI εφαρμογές, όπως και γενικότερα διάφορες cloud εφαρμογές, έχουν την τάση να είναι πιο εύκολες και απλές για τους τελικούς χρήστες όσον αφορά στη λειτουργία και την εγκατάσταση.
- Ταχύτητα ανάπτυξης
- Οι cloud εφαρμογές είναι απλουστευμένες, αφού δεν απαιτούν επιπλέον εγκαταστάσεις υλικού ή λογισμικού.
- Εξέλιξη και ελαστικότητα
- Οι cloud εφαρμογές έχουν την δυνατότητα γρήγορης κλιμάκωσης, έτσι ώστε να καταστεί δυνατή η αύξηση του αριθμού των χρηστών εντός ενός οργανισμού.
- Προσβασιμότητα

Οι Cloud BI εφαρμογές είναι προσβάσιμες από κάθε πρόγραμμα περιήγησης στο διαδίκτυο ή από οποιαδήποτε κινητή συσκευή. Το Cloud BI έχει εξαιρετικές δυνατότητες, όπως το cloud computing. Οι επιχειρήσεις τείνουν να κάνουν μεταφορά των υπηρεσιών τους στο cloud και να απολαμβάνουν τα θετικά αποτελέσματα και τα οφέλη, όπως είναι το χαμηλό κόστος, τα αυξημένα επίπεδα ταχύτητας εγκατάστασης και φυσικά την ευκολία στη χρήση. Οι Cloud εφαρμογές όσον αφορά στην ισχύ τους είναι ισοδύναμες με on-premise εφαρμογές. Οι δυνατότητες που ως σήμερα ήταν σε διαθεσιμότητα μόνο σε τοπικά εγκατεστημένα λογισμικά, τώρα προσφέρονται με μεγάλη ευκολία. Το Cloud BI τείνει διαρκώς να κατακτά εξέχοντα ρόλο στις αναλυτικές επιχειρηματικής ευφυΐας και στις στρατηγικές των επιχειρήσεων, προσφέροντας στους τελικούς χρήστες πρόσβαση σε κρίσιμα για εκείνους δεδομένα, σε πραγματικό χρόνο.

4

Μεθοδολογία

4.1 Τεχνολογία



Εικόνα 4.1: Cloud Data Warehouse & BI

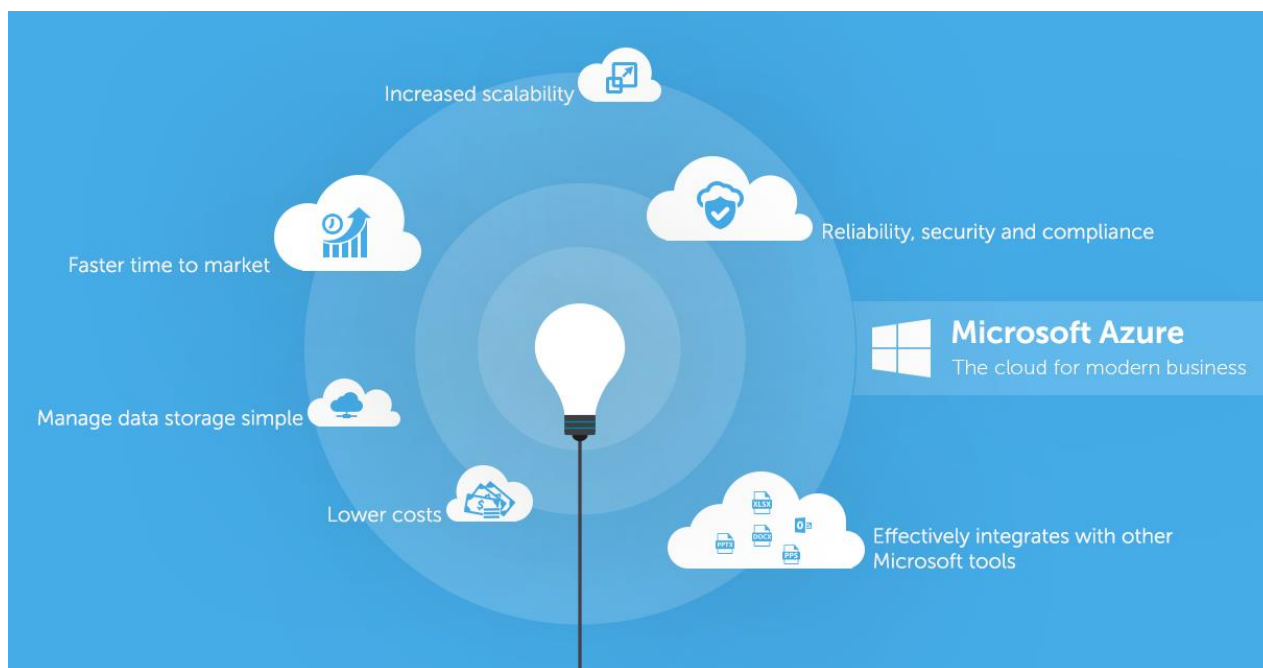
Για την υλοποίηση της διπλωματικής αυτής εργασίας χρησιμοποιήθηκαν Cloud τεχνολογίες, τόσο για την ETLδιαδικασία, την δημιουργία του Data Warehouse καθώς και του Visualization. Η διαδικασία χωρίζεται στα παρακάτω 5 στάδια

- **Εξαγωγή δεδομένων (Extraction of Data):** Σε αυτό το στάδιο πραγματοποιείται η διαδικασία εξαγωγής των δεδομένων από τα αντίστοιχα συστήματα ERP και CRM

- **Ανέβασμα Δεδομένων (Ingest):** Σε αυτό το στάδιο φορτώνουμε τα δεδομένα μας στο Cloud και συγκεκριμένα στο Azure Data Factory που είναι μια ETL Cloud Υπηρεσία της Microsoft. [14]
- **Αποθήκευση (Store):** Αφού έχει ολοκληρωθεί η διαδικασία ETL τα δεδομένα μας αποθηκεύονται στο Azure Data Lake, το οποίο δύναται να υποστηρίξει διαφόρων ειδών δεδομένα και είναι κατάλληλο για Data Warehouses και ανάλυση δεδομένων [15]
- **Προετοιμασία δεδομένων (Prep and Train):** Σε αυτό το στάδιο κανονικά τα δεδομένα χωρίζονται σε δύο κατηγορίες training και testing sets κάτι που βοηθάει πολύ σε περίπτωση Data Mining. Στα πλαίσια της εργασίας αυτής δεν έχει υλοποιηθεί.
- **Μετασχηματισμός και τελικό προϊόν (Model & Serve):** Σε αυτό το στάδιο έχουμε δημιουργήσει ένα Azure SQL Server που έχει την βάση στην οποία θα κάνουμε την ανάλυση. Σε αυτό το σημείο δημιουργούμε κατάλληλα Views στην βάση μας μέσω του Microsoft SQL Server Management Studio για να κατασκευάσουμε το Data Warehouse. Εφαρμόζουμε στο Data Warehouse την ανάλυση και προγνωστική αναλυτική μέσω του Microsoft Power BI και κατασκευάζουμε το τελικό προϊόν, δηλαδή τα διαδραστικά Reports. Τέλος γίνεται το ανέβασμα αυτών των Reports στην Cloud Υπηρεσία το Power BI As a service.

4.2 Σχεδιαστικές Επιλογές (Εργαλεία και Τεχνολογίες)

Σε παγκόσμιο επίπεδο το 90% των εταιρειών που έχουν στραφεί σε Cloud Τεχνολογίες χρησιμοποιούν το Microsoft Azure. Μέσω αυτού μπορούν να δημιουργήσουν, καθώς και να διαχειριστούν από εύκολες έως και αρκετά περίπλοκες καθημερινές τους ανάγκες καθώς υποστηρίζει ένα μεγάλο εύρος γλωσσών προγραμματισμού, λειτουργικών συστημάτων βάσεων δεδομένων και συσκευών ενώ ταυτόχρονα παρέχει και την αξιοπιστία της Microsoft [16]. Ακολουθώντας τις τάσεις της αγοράς για την υλοποίηση μας τόσο για την ETL διαδικασία αλλά και αποθήκευση των δεδομένων μας χρησιμοποιήσαμε την Cloud Υπηρεσία του Microsoft (Azure).



Εικόνα 4.2: Microsoft Azure

Κάποια από τα πλεονεκτήματα που μας δίνει είναι τα εξής [16]:

- Εύκολη Διαχείριση των δεδομένων (ETL)
- Χαμηλότερα κόστη σε σχέση με άλλες Cloud υπηρεσίες
- Ασφάλεια των δεδομένων μας από φυσικές καταστροφές και διαδικτυακές επιθέσεις
- Αρίστη συνεργασία με άλλα προϊόντα της Microsoft
- Δυνατότητα Επεκτασιμότητας
- Ενσωματωμένη δυνατότητα αναλυτικής και Business Intelligence



Εικόνα 4.3: Sql Server

Για την κατασκευή του Data Warehouse χρησιμοποιήθηκε το SQL Server Management Studio 2018. Καθώς πέρα από το γεγονός ότι είναι και αυτό ένα προϊόν της Microsoft επομένως συνεργάζεται καλύτερα με τα υπόλοιπα

εργαλεία που έχουμε χρησιμοποιήσει μας δίνει την δυνατότητα εύκολα να δούμε τα δεδομένα που

έχουμε ανεβασμένα στο Microsoft Azure καθώς και να δημιουργήσουμε οποιαδήποτε query είναι απαραίτητα για την υλοποίηση μας.



Εικόνα 4.4: Power BI

Τέλος για την ανάλυση μας καθώς και την κατασκευή των τελικών αναφορών που έχουν σαν αποδέκτη των πελάτη χρησιμοποιήθηκε το Microsoft Power BI. Τοποίο από το 2019 και μετά έχει αρχίσει να αποκτά μεγαλύτερο μερίδιο στην αγορά χάρη στις παρακάτω δυνατότητες που παρέχει:

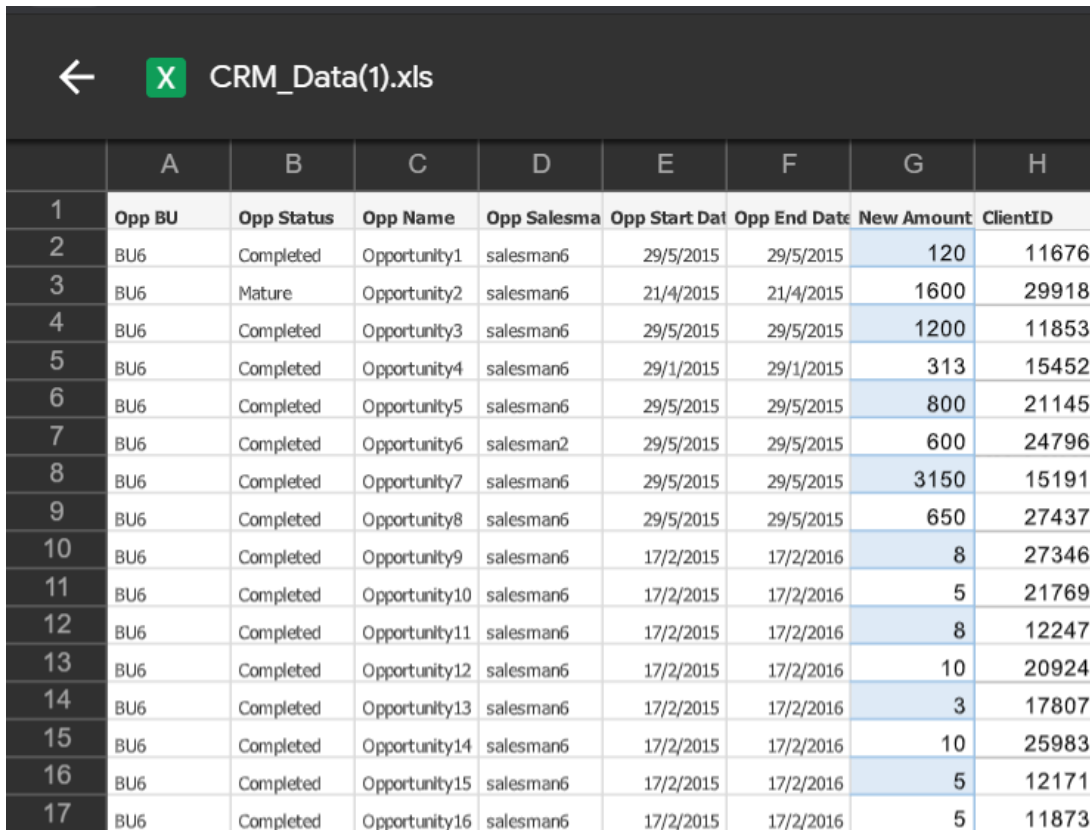
- Συμβατότητα με την οικογένεια τεχνολογιών Microsoft
- Τεράστια επιλογή data sources (SQL Server, Excel, CSVs, XMLs, Oracle DBs, MySQL, κ.ά.)
- Σύνδεση με Microsoft Azure και άλλες πλατφόρμες Cloud
- Εύκολη διανομή dashboard (Power BI as a service)
- Δυνατότητα στατιστικής ανάλυσης με χρήση γλώσσας R
- Build-in αλγορίθμους προγνωστικής αναλυτικής
- Διαδραστικά και όμορφα γραφικά
- Δυνατότητα Mobile υπηρεσίας
- Δυνατότητα Data Mining

4.3 ERP, CRM εισαγωγή δεδομένων

Για την υλοποίηση μας χρησιμοποιήθηκαν δεδομένα από δύο διαφορετικές πηγές. Από ένα ERP σύστημα και ένα CRM και απεικονίζουν οικονομικά και πελατειακά στοιχεία μιας εταιρείας μεταπώλησης. Σε αυτό το σημείο να τονίσουμε ότι τα δεδομένα μας έχουν υποστεί επεξεργασία

με χρήση κατάλληλων αλγορίθμων μετά την εξαγωγή τους από τα αντίστοιχα συστήματα και σε καμία περίπτωση δεν είναι τα πραγματικά.

Τα πελατειακά στοιχεία εξήχθησαν από το σύστημα σε μία flat μορφή αρχείου excel την μορφή της οποίας μπορούμε να δούμε στην εικόνα 4-5.

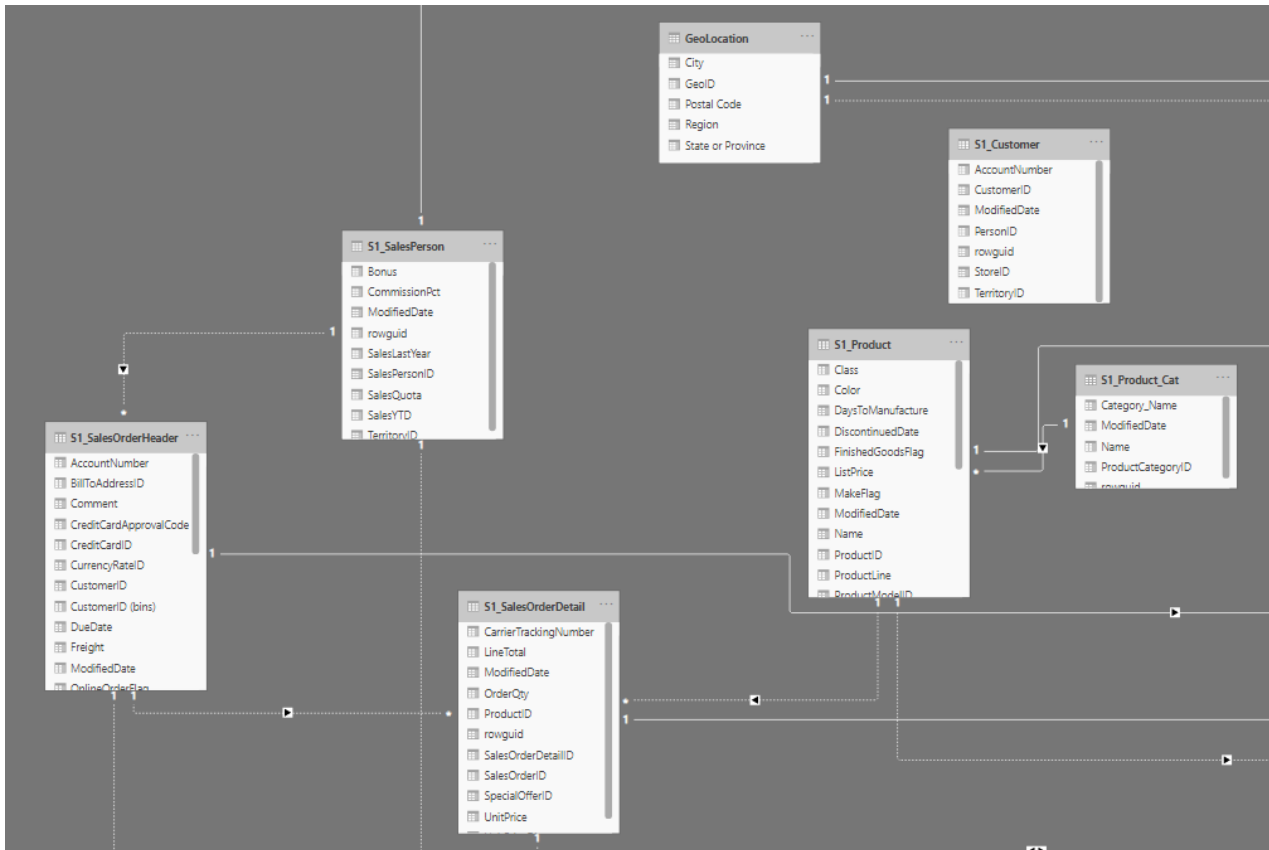


	A	B	C	D	E	F	G	H
	Opp BU	Opp Status	Opp Name	Opp Salesman	Opp Start Date	Opp End Date	New Amount	ClientID
1	BU6	Completed	Opportunity1	salesman6	29/5/2015	29/5/2015	120	11676
2	BU6	Mature	Opportunity2	salesman6	21/4/2015	21/4/2015	1600	29918
3	BU6	Completed	Opportunity3	salesman6	29/5/2015	29/5/2015	1200	11853
4	BU6	Completed	Opportunity4	salesman6	29/1/2015	29/1/2015	313	15452
5	BU6	Completed	Opportunity5	salesman6	29/5/2015	29/5/2015	800	21145
6	BU6	Completed	Opportunity6	salesman2	29/5/2015	29/5/2015	600	24796
7	BU6	Completed	Opportunity7	salesman6	29/5/2015	29/5/2015	3150	15191
8	BU6	Completed	Opportunity8	salesman6	29/5/2015	29/5/2015	650	27437
9	BU6	Completed	Opportunity9	salesman6	17/2/2015	17/2/2016	8	27346
10	BU6	Completed	Opportunity10	salesman6	17/2/2015	17/2/2016	5	21769
11	BU6	Completed	Opportunity11	salesman6	17/2/2015	17/2/2016	8	12247
12	BU6	Completed	Opportunity12	salesman6	17/2/2015	17/2/2016	10	20924
13	BU6	Completed	Opportunity13	salesman6	17/2/2015	17/2/2016	3	17807
14	BU6	Completed	Opportunity14	salesman6	17/2/2015	17/2/2016	10	25983
15	BU6	Completed	Opportunity15	salesman6	17/2/2015	17/2/2016	5	12171
16	BU6	Completed	Opportunity16	salesman6	17/2/2015	17/2/2016	5	11873

Εικόνα 4.5: Δεδομένα CRM

Σε αυτό το αρχείο παρουσιάζονται δεδομένα σε σχέση με ευκαιρίες που είχε η επιχείρησή μας να κάνει μια πώληση (Opp BU) την τρέχουσα κατάσταση στην οποία βρίσκεται η ευκαιρία (Opp Status). Ακολουθούν πληροφορίες σχετικά με την ημερομηνία έναρξης και ολοκλήρωσης της ευκαιρίας (Opp Start Date / Opp End Date), το όνομα της ευκαιρίας (Opp name),σε ποιόν πωλητή ανήκει (Opp Salesman) καθώς και τι κέρδος θα φέρει.

Τα δεδομένα επιχειρησιακών πόρων έχουν εξαχθεί από το ERP σύστημα επιλέγοντας το χρονικό διάστημα για το οποίο θέλουμε δεδομένα και επιλέγουμε να εξάγουμε μια σχεσιακή βάση δεδομένων σε μορφή αρχείου (dacrps file) η οποία περιλαμβάνει τόσο τα δεδομένα όσο και το σχήμα της βάσης και είναι μορφή που υποστηρίζεται από τον SQL Server [17]



Εικόνα 4.6: Δεδομένα ERP

Μπορούμε να δούμε ότι αποτελείται από 6 πίνακες οι οποίοι συνδέονται μεταξύ τους μέσω foreign keys. Βλέπουμε ότι έχουμε την διάσταση του πελάτη, του προϊόντος, της τοποθεσίας, της κατηγορίας προϊόντων καθώς και πίνακες οι οποίοι περιέχουν πιο σύνθετη πληροφορία η οποία αφορά παραγγελίες και οικονομικά στοιχεία τα οποία θα δούμε και πιο αναλυτικά στην συνέχεια. Σε αυτό το σημείο κατανοούμε το business της επιχείρησης πως δηλαδή αυτή λειτουργεί και αποφασίζουμε ποιους δείκτες και ποια στοιχεία η επιχείρηση θα προτιμούσε να παρακολουθεί.

4.4 Επεξεργασία Δεδομένων στο Cloud

Εφόσον έχουμε πλέον τα δεδομένα της ανάλυσης μας σειρά έχει να τα ανεβάσουμε στο Microsoft Azure Cloud και να κάνουμε όλους τους απαραίτητους μετασχηματισμούς για να τα φέρουμε στην μορφή που χρειάζεται για την ανάλυση μας.

1. Αρχικά πρέπει να φτιάξουμε ένα νέο λογαριασμό στο Microsoft Azure Cloud. Για να το κάνουμε αυτό πληκτρολογούμε την διεύθυνση <https://azure.microsoft.com/en-us/free/> και συμπληρώνουμε τα στοιχεία που εμφανίζονται στην εικόνα 4-7. Ουσιαστικά με το free plan της Microsoft μας δίνει δωρεάν 150\$ σε credits τα οποία αντιστοιχούν σε πόρους που μπορούμε να

αγοράσουμε μέσα στην Cloudυπηρεσία και διαρκούν είτε μέχρι να τα ξοδέψουμε είτε μέχρι την λήξη της δωρεάν δοκιμής (30 ημερών). Ακόμα μας δίνει κάποιες free υπηρεσίες τις οποίες μπορούμε να χρησιμοποιήσουμε για 12 μήνες από την δημιουργία του λογαριασμού μας.

Microsoft Azure

Sign out

Try Azure for free

Follow these steps to get started. We ask for these details to protect your account and information. There are no upfront charges or fees.

1 About you

Country/Region ⓘ
United Kingdom

Choose the location that matches your billing address. **You cannot change this selection later.** If your country is not listed, the offer is not available in your region. [Learn More](#)

First name

Last name

Email address ⓘ

Phone
Example: 1632 456789

Company VatID ⓘ
Optional

Next

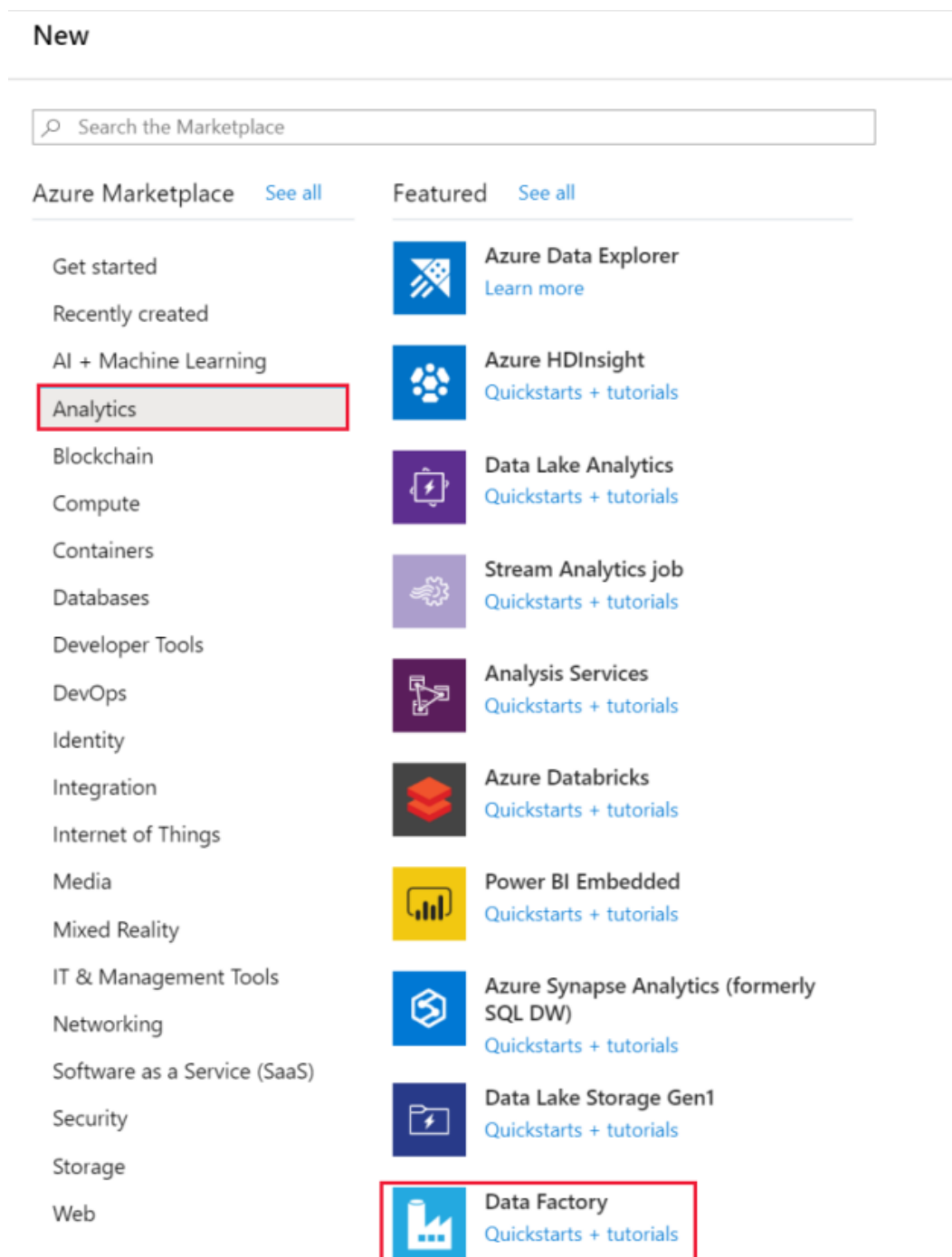
What's included

- ✓ **12 months of free products**
Get free access to popular products like *virtual machines*, *storage*, and *databases* in your first 30 days, and for 12 months after you upgrade your account to pay-as-you-go pricing.
- ✓ **£150 credit**
Use your £150 credit to experiment with any Azure service in your first 30 days—beyond the free product amounts.
- ✓ **25+ always-free products**
Take advantage of more than 25 products, including *serverless*, *containers*, and *artificial intelligence*, that are always free. Get these in your first 30 days, and always—once you choose to upgrade.
- ✓ **No automatic charges**
You won't be charged unless you choose to upgrade. Before the end of your first 30 days, you'll be notified and have the chance to upgrade and start paying only for the resources you use beyond the free amounts.

Chat with Sales

Εικόνα 4.7: Azure Sign Up

2. Αφού έχουμε πλέον έτοιμο τον λογαριασμό μας, σειρά έχει να φτιάξουμε ένα νέο resource Azure Data Factory, για να το κάνουμε αυτό πηγαίνουμε από το Azure Portal στο menu και επιλέγουμε την υπηρεσία Data Factory.



Εικόνα 4.8: Azure Data Factory

3. Δημιουργούμε το Data Factory δίνοντας ένα όνομα που θέλουμε και επιλέγοντας το Subscription που έχουμε δημιουργήσει, δηλαδή το free trial. Επιλέγουμε τοποθεσία, resource group και στη συνέχεια πατάμε “Create”

Home > Data factories > New data factory

New data factory

* Name ⓘ
NewDataFactory20 ✓

* Subscription
Free Trial ▾

* Resource Group ⓘ
☐ Create new ☒ Use existing
 New ▾

Version ⓘ
V2 ▾

* Location ⓘ
North Europe ▾

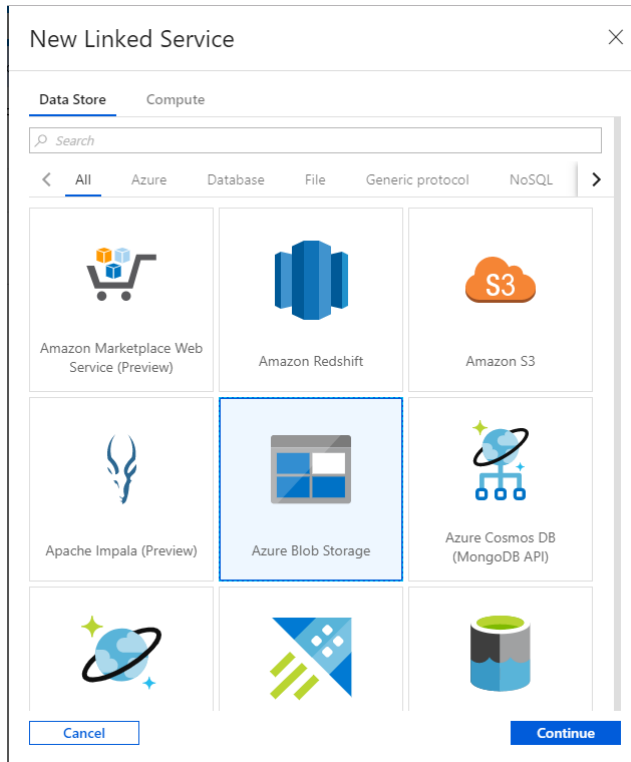
Integrate with GIT source control (Azure DevOps GIT or GitHub) to do collaboration, source control, change tracking, change difference, continuous integration and deployment etc

☐ Enable GIT ⓘ

[Create](#) [Automation options](#)

Εικόνα 4.9: Creation of Azure Data Factory

4. Σειρά έχει να δημιουργήσουμε ένα Blob Storage το οποίο θα χρησιμοποιεί το Data Factory για αποθήκευση. Το δημιουργούμε ακολουθώντας τα βήματα που βλέπουμε στις παρακάτω εικόνες.



Εικόνα 4.10: Azure Storage

Create storage account

Project details
Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

* Subscription: Free Trial
* Resource group: New
[Create new](#)

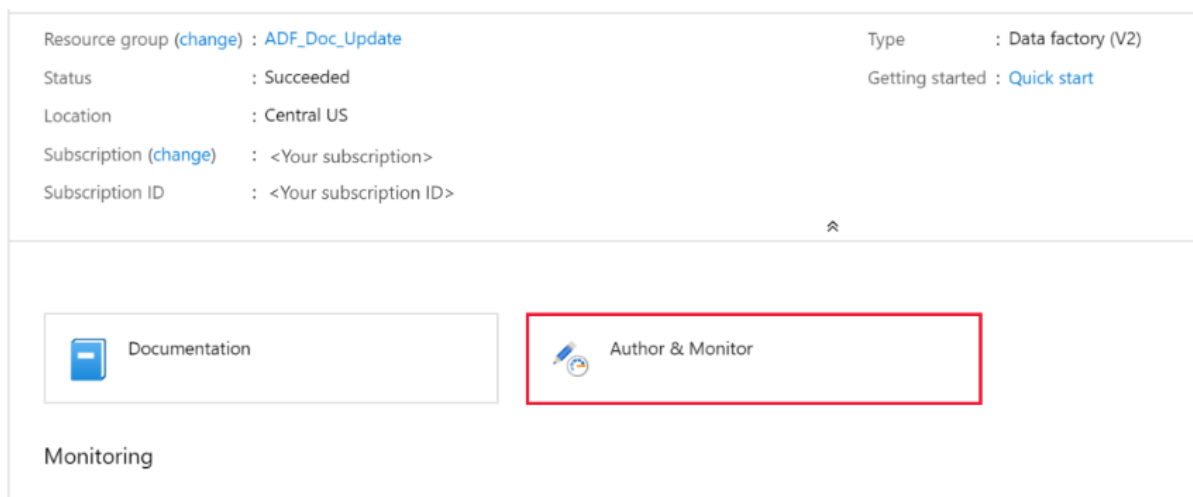
Instance details
The default deployment model is Resource Manager, which supports the latest Azure features. You may choose to deploy using the classic deployment model instead. [Choose classic deployment model](#)

* Storage account name: st2020orage ✓
* Location: (Europe) North Europe
Performance: ☒ Standard ☐ Premium
Account kind: StorageV2 (general purpose v2)
Replication: Read-access geo-redundant storage (RA-GRS)
Access tier (default): ☒ Cool ☐ Hot

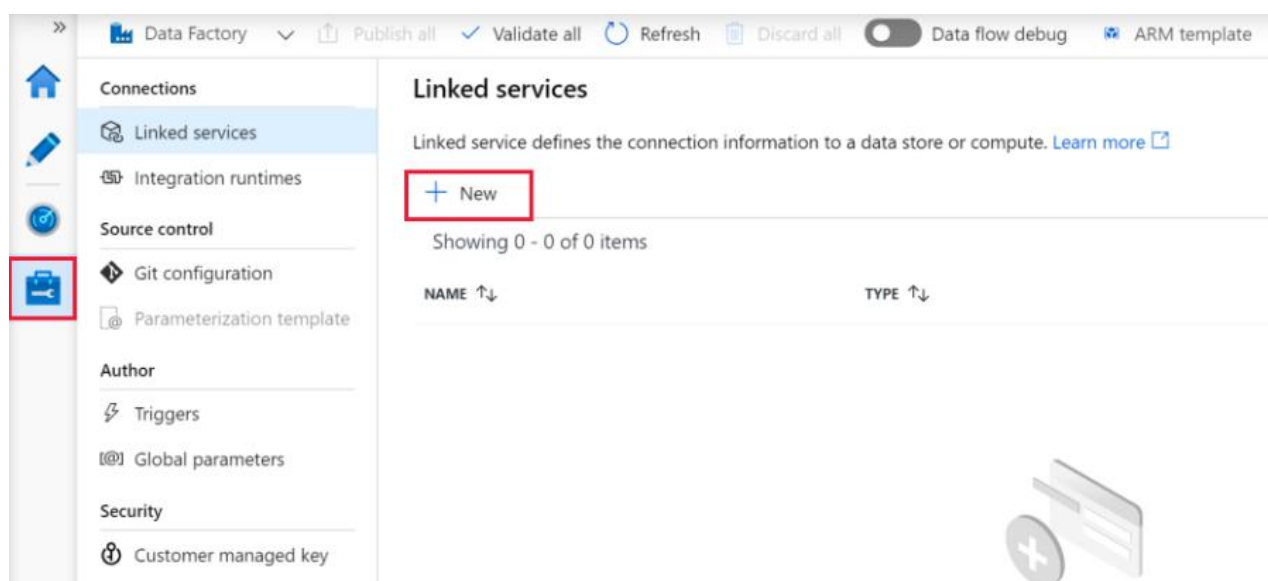
[Review + create](#) [< Previous](#) [Next : Advanced >](#)

Εικόνα 4.11: Creation of Azure Storage

5. Τώρα πρέπει να συνδέσουμε το Data Factory με το Storage. Για να γίνει αυτό πηγαίνουμε στο Data Factory και επιλέγουμε Author and Monitor και στην σελίδα που ανοίγει συμπληρώνουμε ότι ζητείτε και προσθέτουμε τον storageχώρο που έχουμε δημιουργήσει.



Εικόνα 4.12: Data Factory (Author & Monitor)



Εικόνα 4.13: Data Factory Link Service

New linked service (Azure Blob Storage)

Name *
AzureStorageLinkedService

Description

Connect via integration runtime *
AutoResolveIntegrationRuntime

Authentication method
Account key

Connection string Azure Key Vault

Account selection method
☒ From Azure subscription ☐ Enter manually

Azure subscription
<select your Azure subscription here>

Storage account name *
<select your Storage account name here>

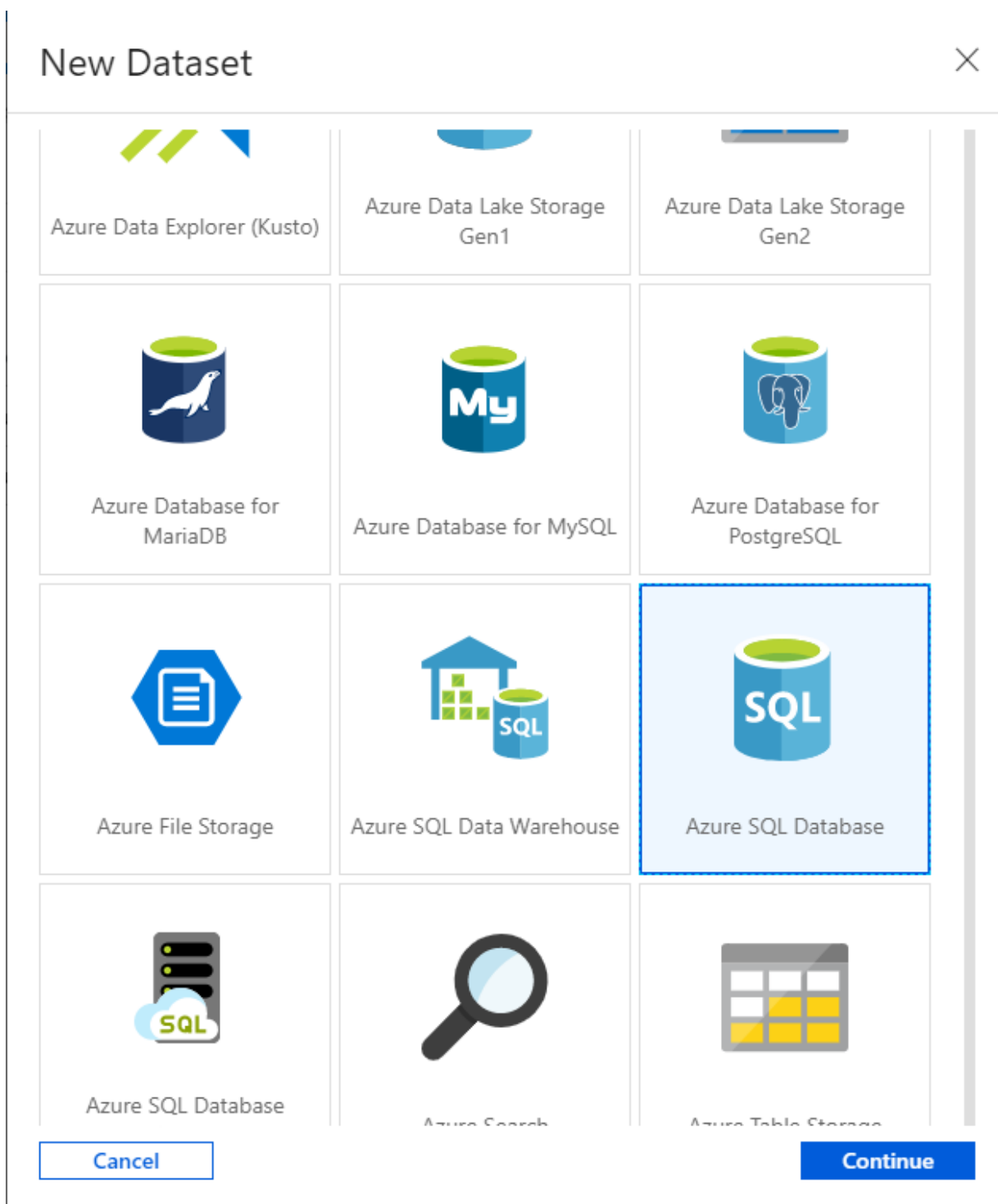
Additional connection properties
+ New

Connection successful

Create Back Test connection Cancel

Εικόνα 4.14: New Linked Service

6. Επόμενο βήμα είναι να φτιαχτούν τα Data Sets που θα φορτωθούν στο Data Factory. Σαν Data Set επιλέξαμε να φτιάξουμε δύο βάσεις δεδομένων μια για κάθε τύπο δεδομένων. Επομένως πρέπει να φτιάξουμε και έναν Sql Server για την διαχείριση τους.[18]



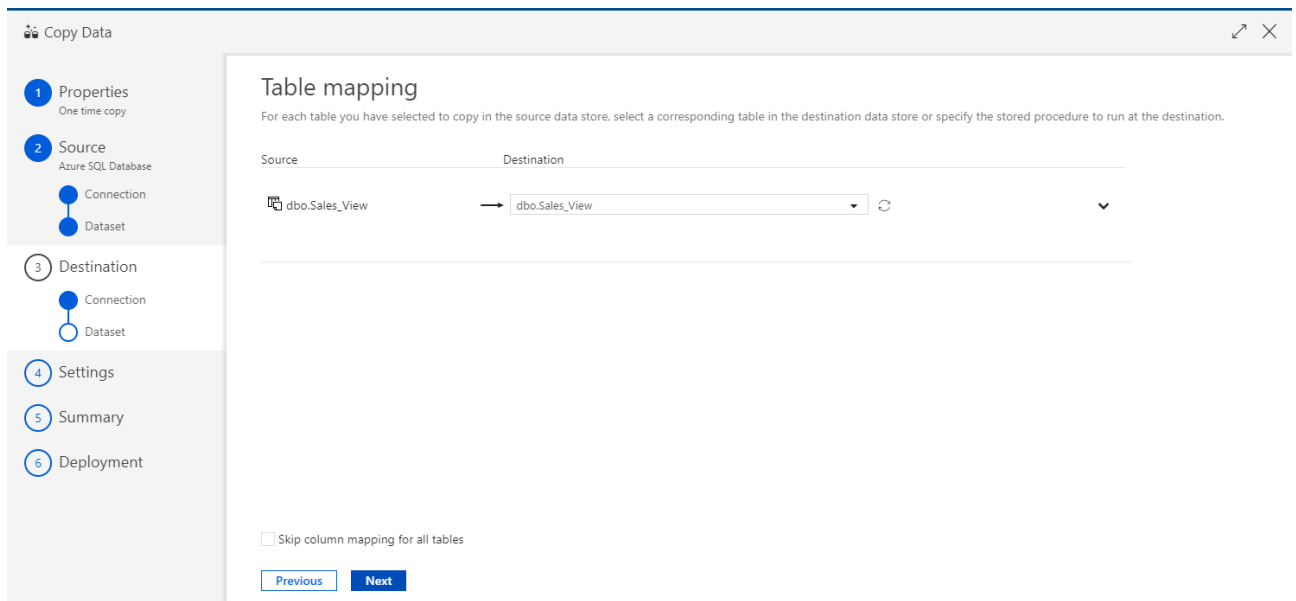
Εικόνα 4.15: New Data Set

Εικόνα 4.16: Δημιουργία Sql Server

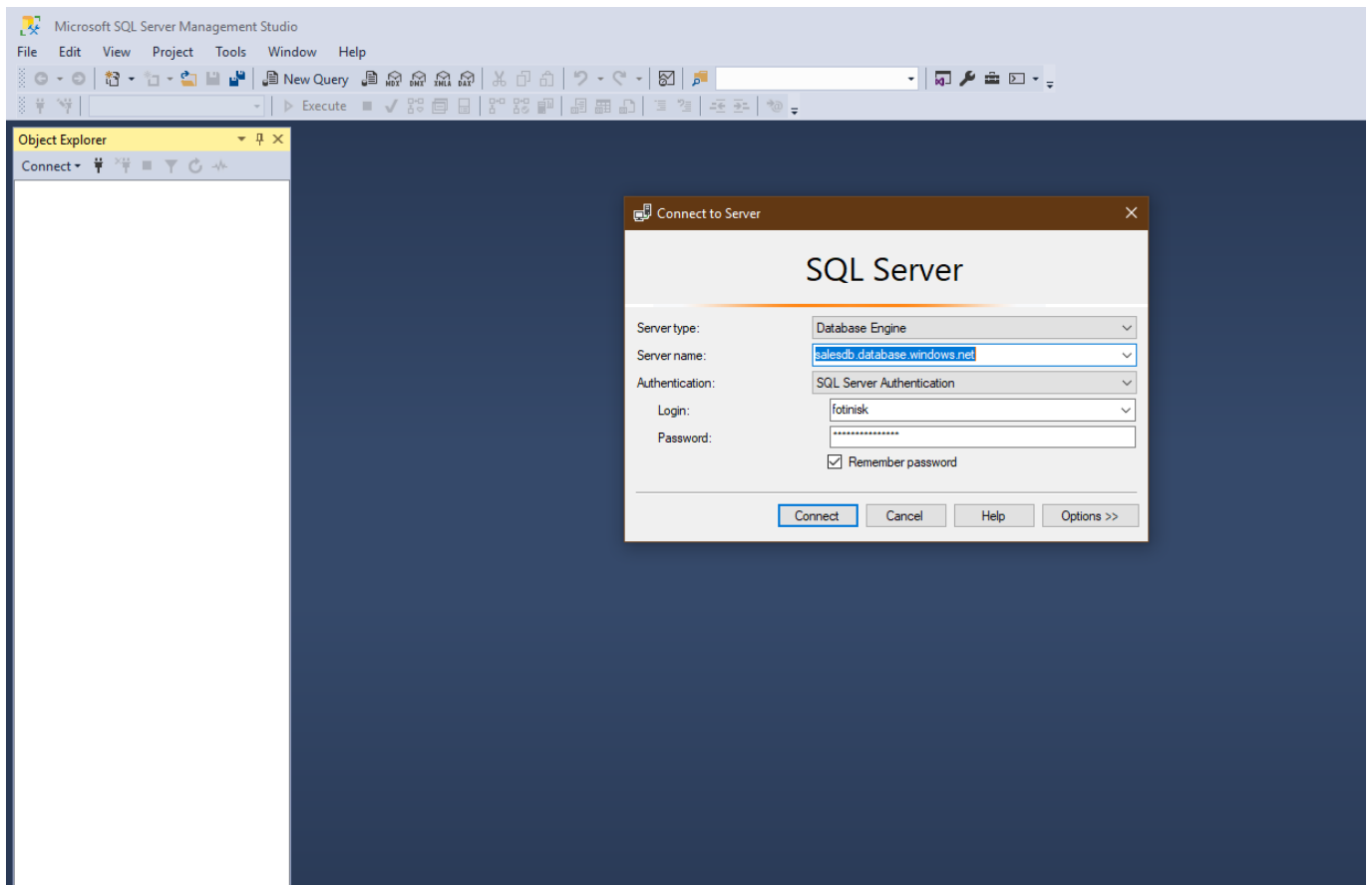
- Εφόσον έχει κατασκευαστεί η βάση δεδομένων, σειρά έχει το migration των δεδομένων στο Cloud. Για να γίνει αυτό αρχικά πρέπει να γίνει login στη βάση δεδομένων που δημιουργήθηκε προηγουμένως και μέσω του Query Editor να γίνει import του dacpac αρχείου, δηλαδή των στοιχείων του ERP συστήματος. Αφού πατηθεί το OK για να γίνει το import πρέπει για κάθε πίνακα του σχήματος να επιλεγεί ο Source καθώς και ο Destination πίνακας. Ακολουθώντας την ίδια διαδικασία φτιάχτηκε και μια βάση δεδομένων (CRM) για τα CRM δεδομένα.

Εικόνα 4.17: Import Database

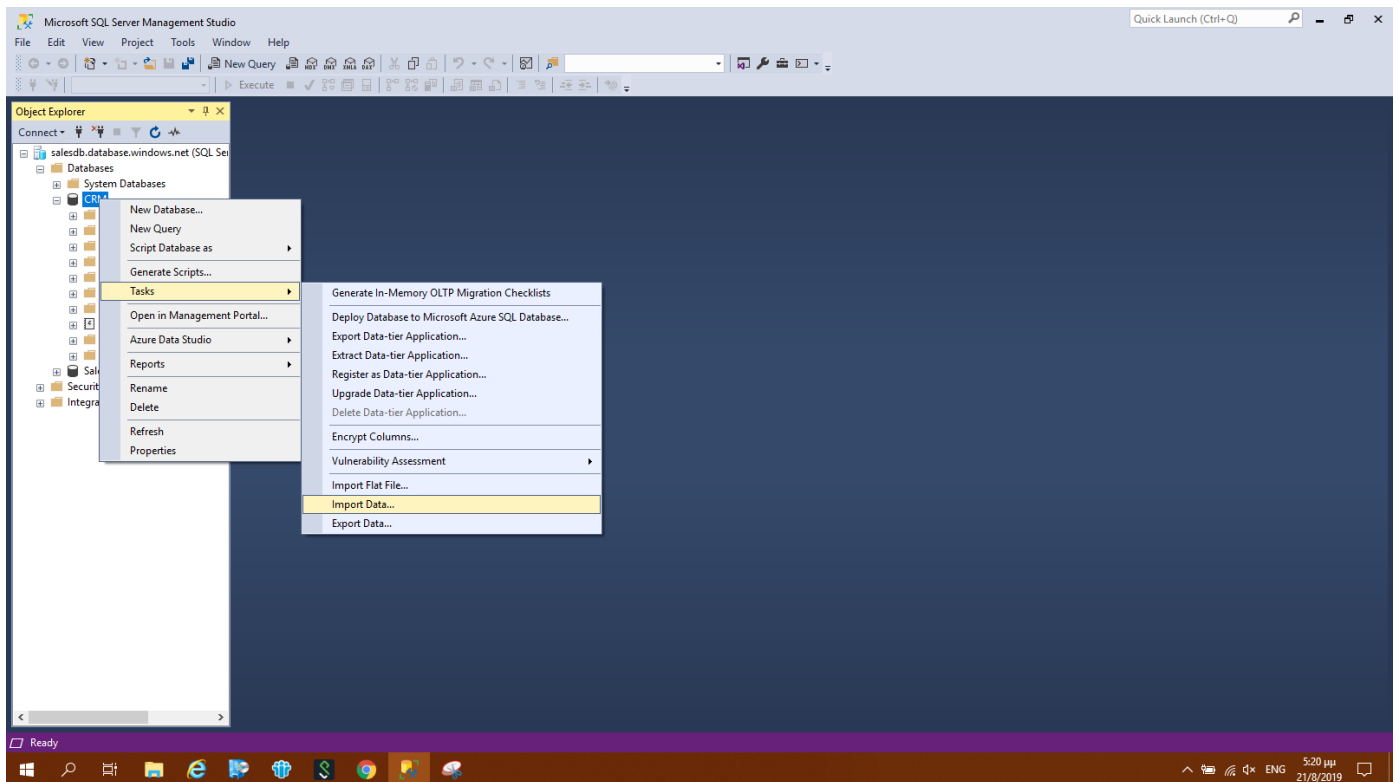
8. Σειρά έχει η μεταφόρτωση του excel αρχείου στην βάση δεδομένων CRM. Για αυτή την διαδικασία θα ακολουθηθεί διαφορετική διαδικασία από αυτή που δείξαμε. Για την εργασία αυτή θα χρησιμοποιηθεί το εργαλείο SSMS. Αρχικά θα γίνει η σύνδεση στον Server που έχει φτιαχτεί στο Azure. Από το import export wizard πρέπει να επιλεγεί σαν source file το CRM και σαν destination που η βάση που φτιάχτηκε στο Azure. Στη συνέχεια πρέπει να επιλεγθούν οι πίνακες οι οποίοι θέλουμε να είναι στην τελική βάση δεδομένων.



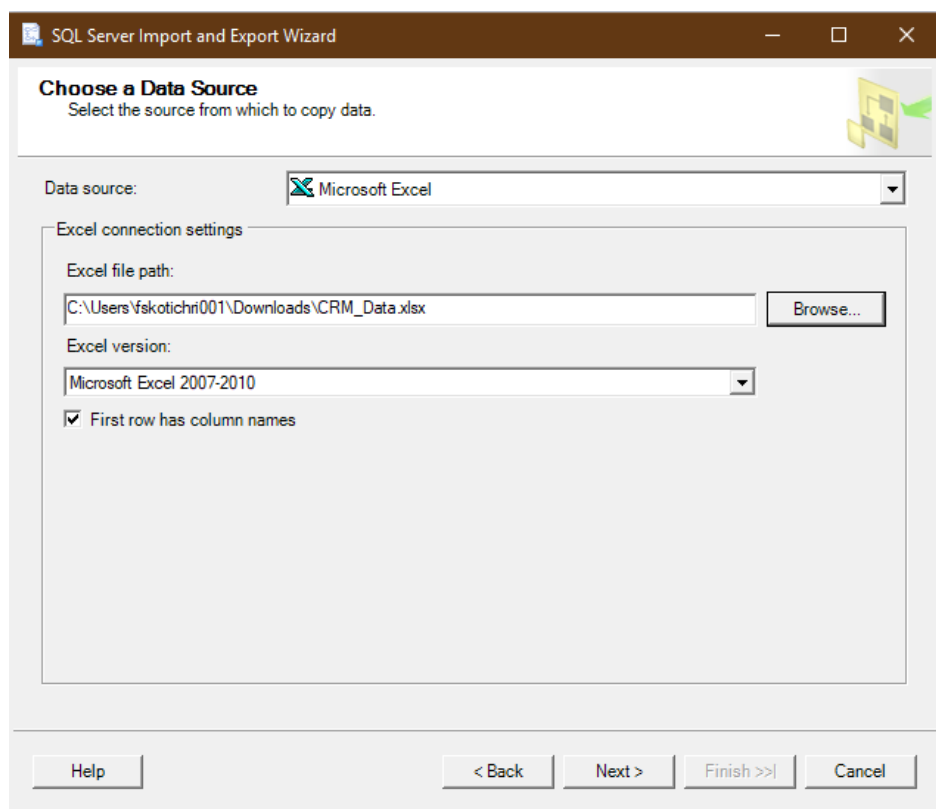
Εικόνα 4.18: Table Mapping



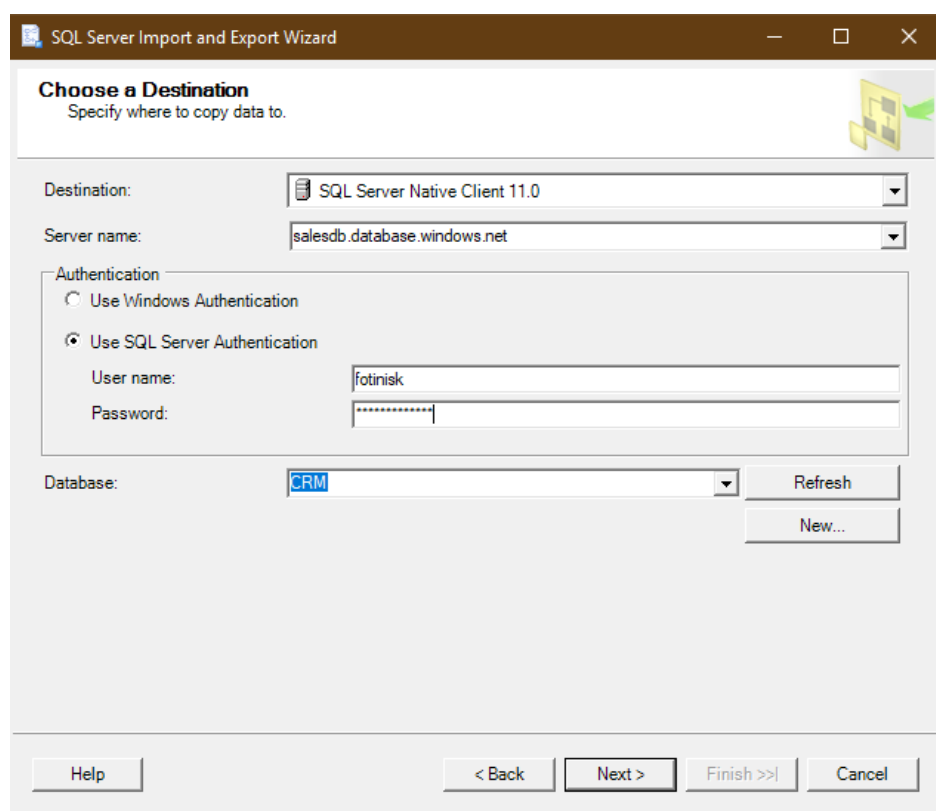
Εικόνα 4.19: SSMS Connection to Sql Server



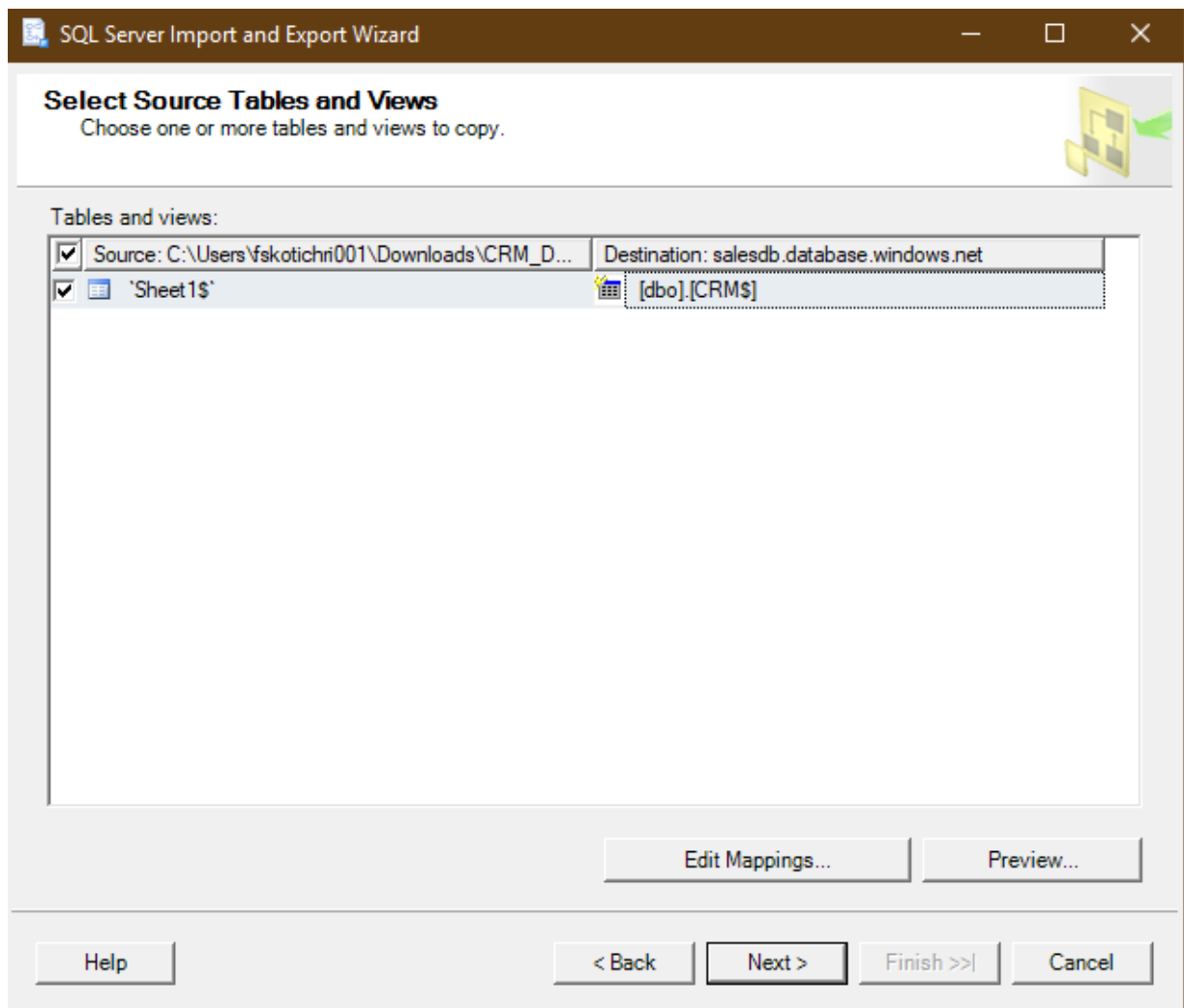
Εικόνα 4.20: SSMS Import File



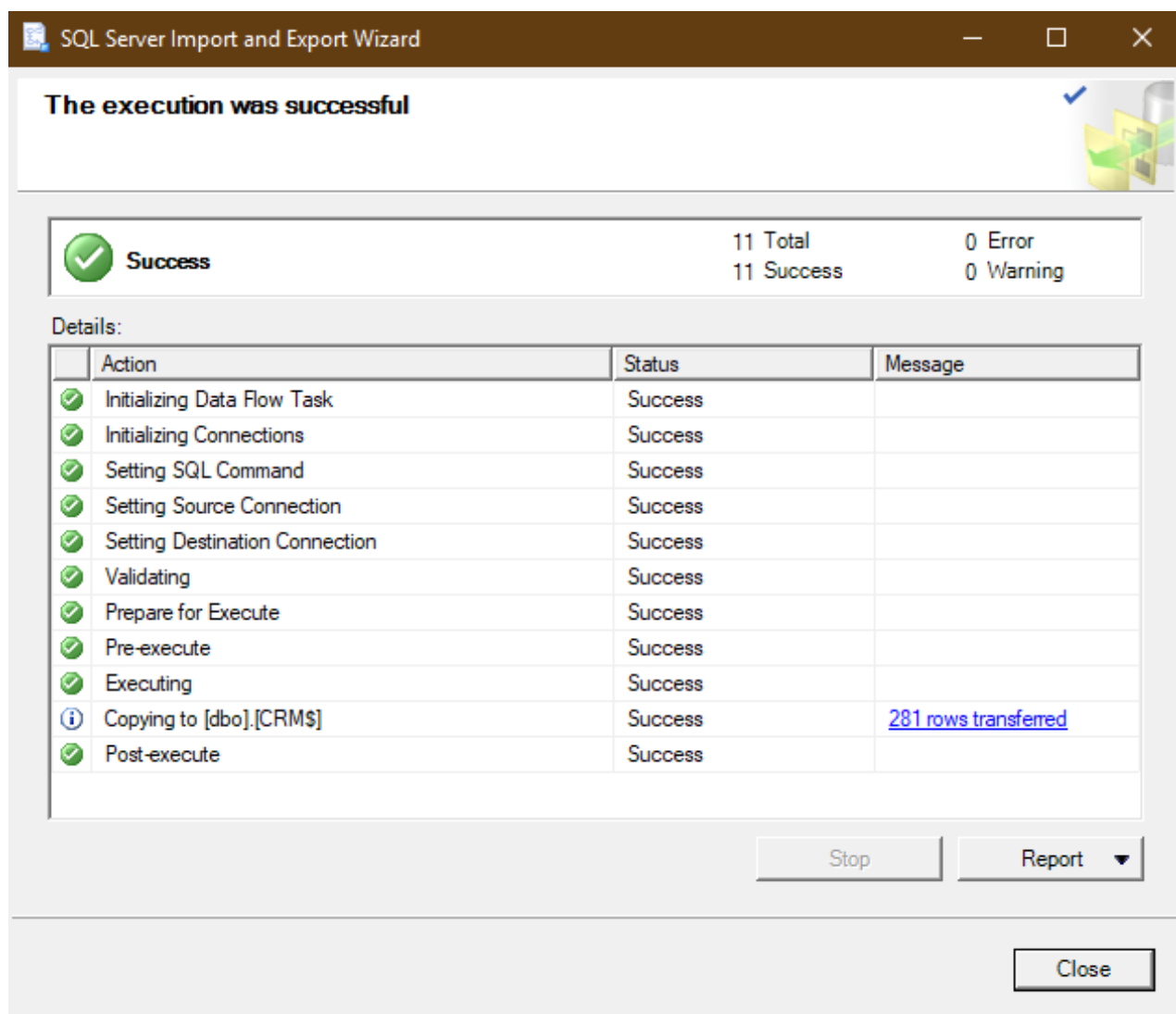
Εικόνα 4.21: Choose Data Source



Εικόνα 4.22: Select Destination

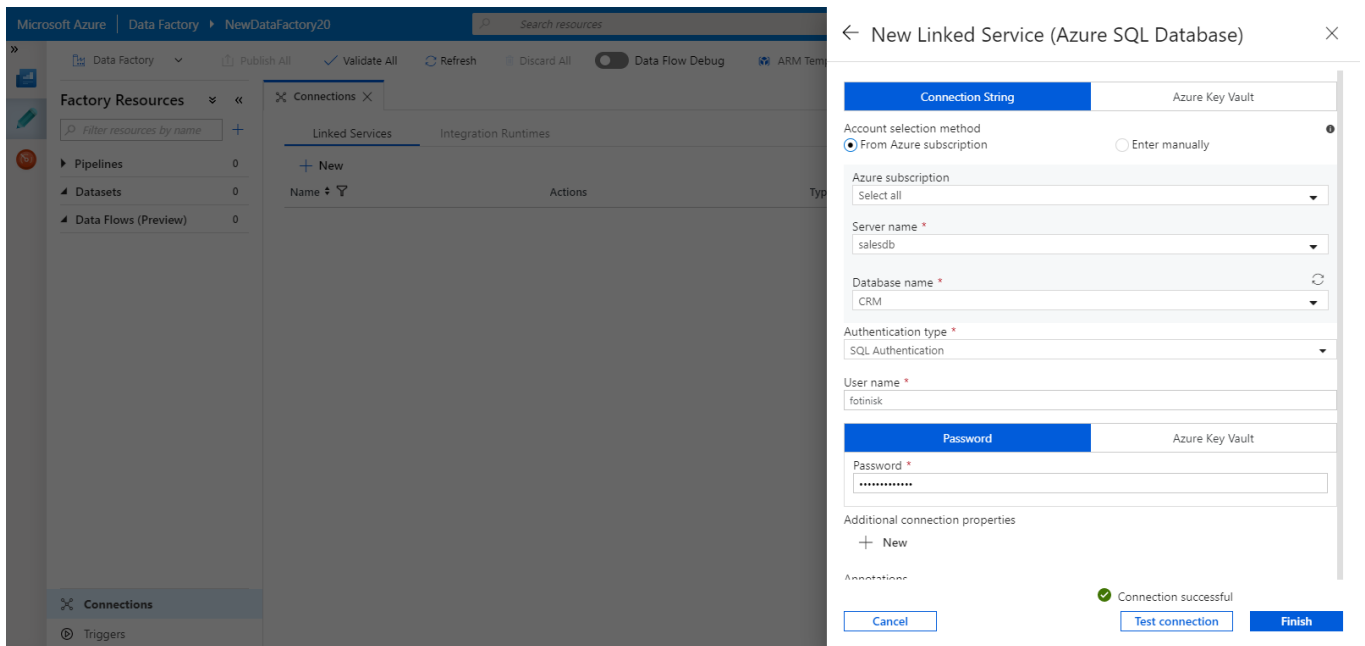


Εικόνα 4.23: Select Source Tables and Views



Εικόνα 4.24: Import of CRM into Sql Server

- Αφού έχουμε κατασκευάσει τις βάσεις δεδομένων πρέπει να δημιουργήσουμε μια διασύνδεση μεταξύ του Data Factory και των βάσεων μας. Συνολικά δηλαδή θα φτιάξουμε δύο συνδέσεις.



Εικόνα 4.25: Linked Service CRM

10. Σε αυτό το στάδιο θα δημιουργήσουμε μία νέα βάση δεδομένων στην οποία θα ενοποιήσουμε τα δεδομένα που έχουμε συλλέξει από τις άλλες δύο και στην οποία θα στηριχτεί το Data Warehouse που θα φτιάξουμε στην συνέχεια και επιπλέον θα κατασκευάσουμε και ένα linked Service για την νέα βάση αυτή και το Data Factory.

Home > SQL databases > Create SQL Database

Create SQL Database

Microsoft

Warning: Changing basic options may reset selections you have made. Please review all options prior to creating the database.

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

* Subscription: Free Trial

* Resource group: New

database details

Enter required settings for this database, including picking a logical server and configuring the compute and storage resources

* Database name: Finaldb

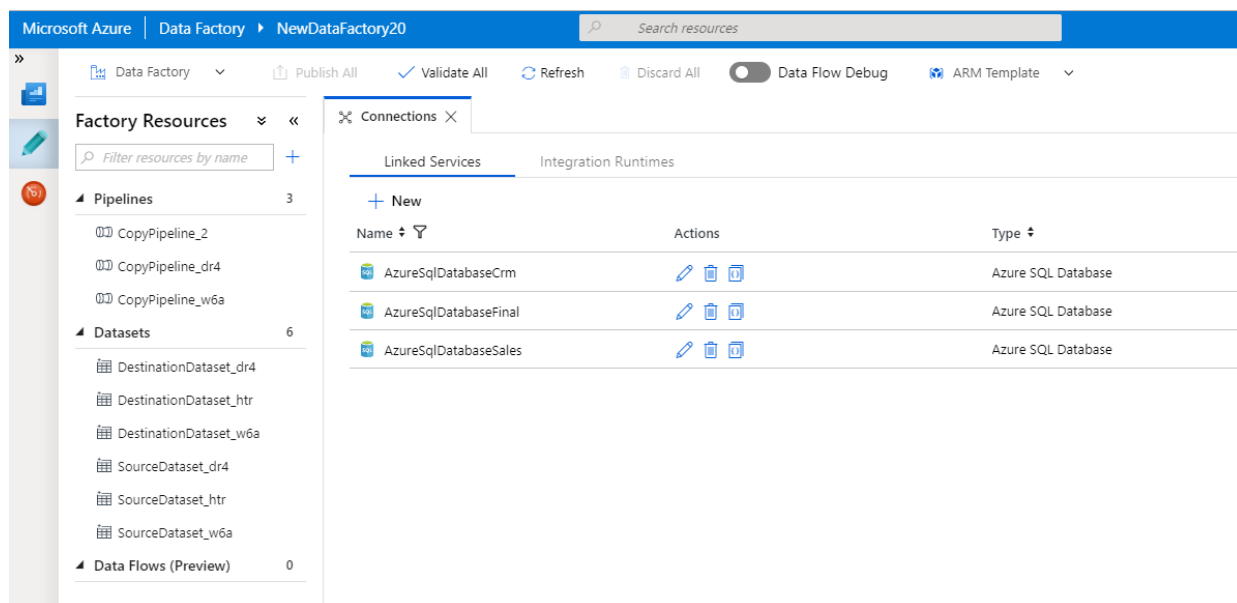
* Server: salesdb (West Europe)

* Want to use SQL elastic pool?: Yes

* Elastic pool: saleselastic (Pricing tier : GeneralPurpose)

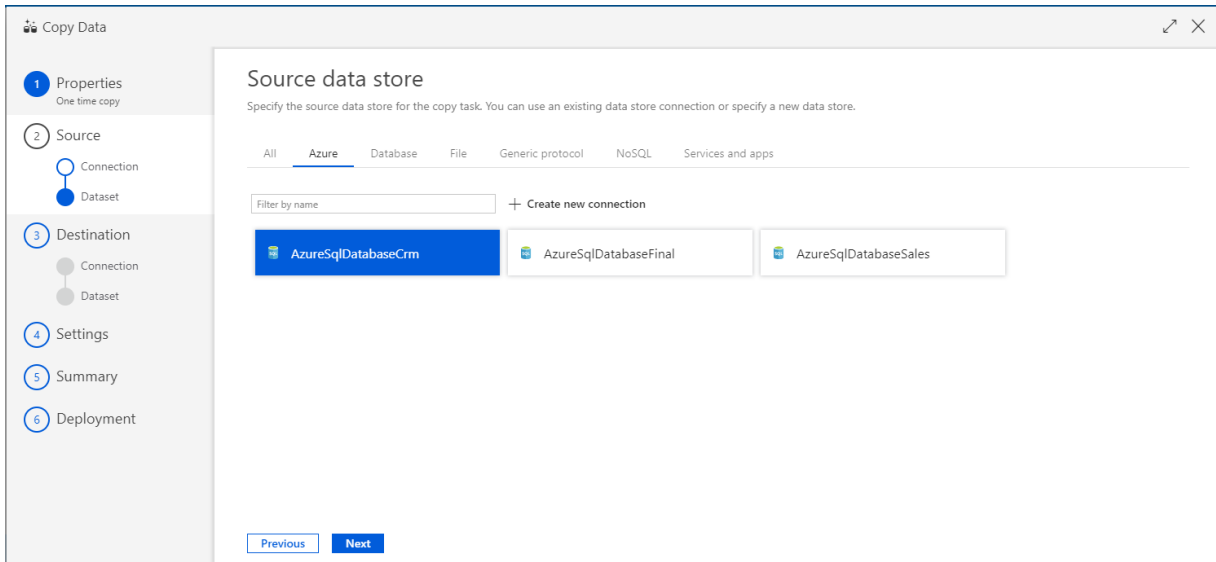
Buttons: Review + create, Next : Additional settings >, Download a template for automation

Εικόνα 4.26: Κατασκευή Τελικής Βάσης

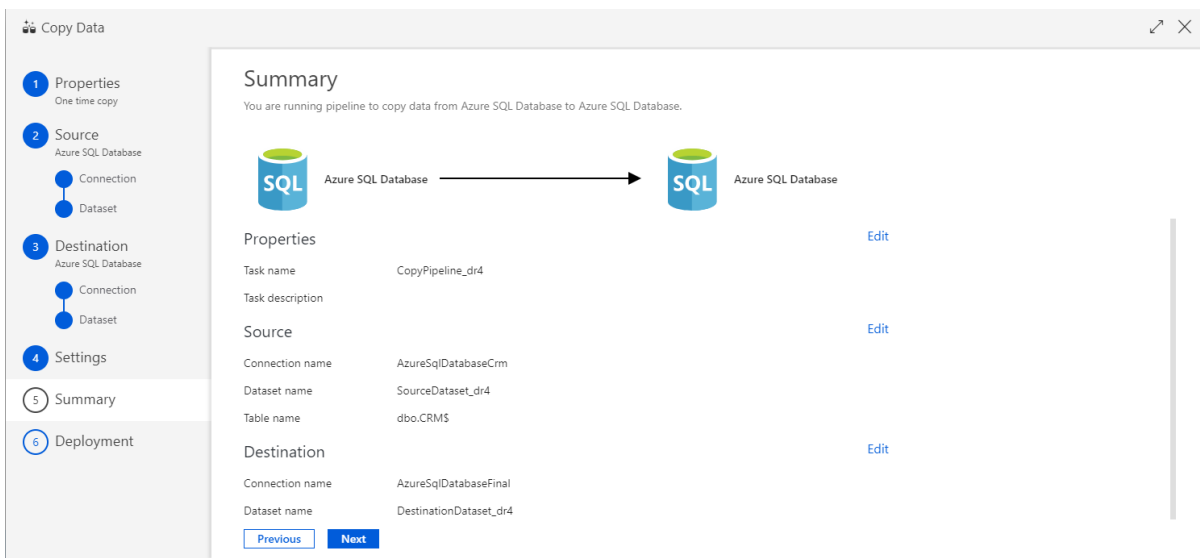


Εικόνα 4.27: Final Linked Services

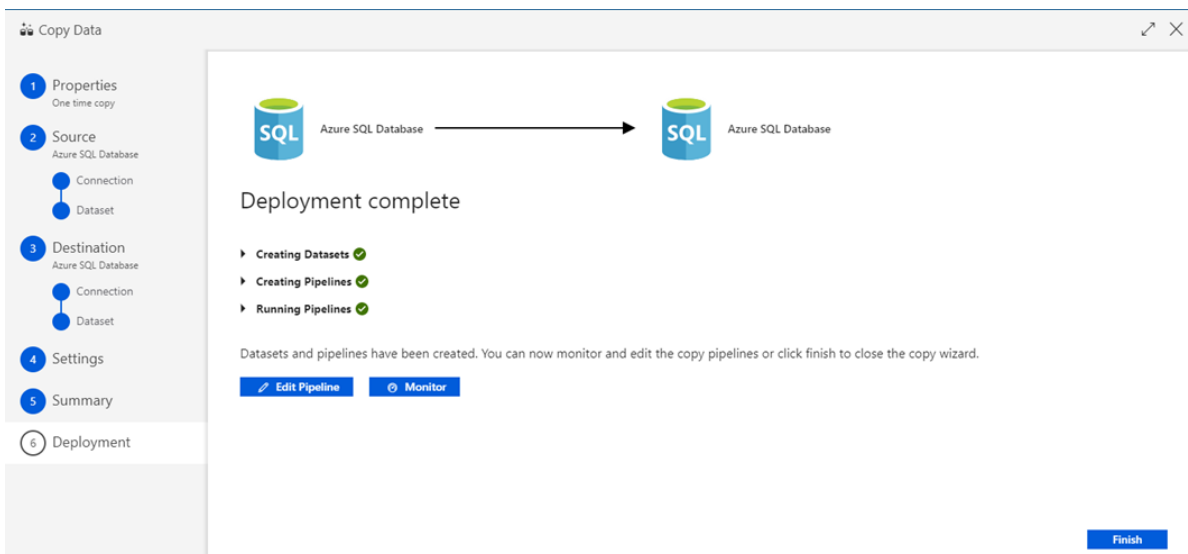
11. Δημιουργούμε δύο pipeline όπως φαίνονται και στην εικόνα παραπάνω. Τα pipeline χρησιμοποιούνται για να αυτοματοποιούν διαδικασίες και μπορεί να περιέχουν ένα ή περισσότερα task. Εμείς θα τα χρησιμοποιήσουμε για να κατασκευάσουμε την τελική δομή της βάσης.



Εικόνα 4.28: Pipeline for CRM



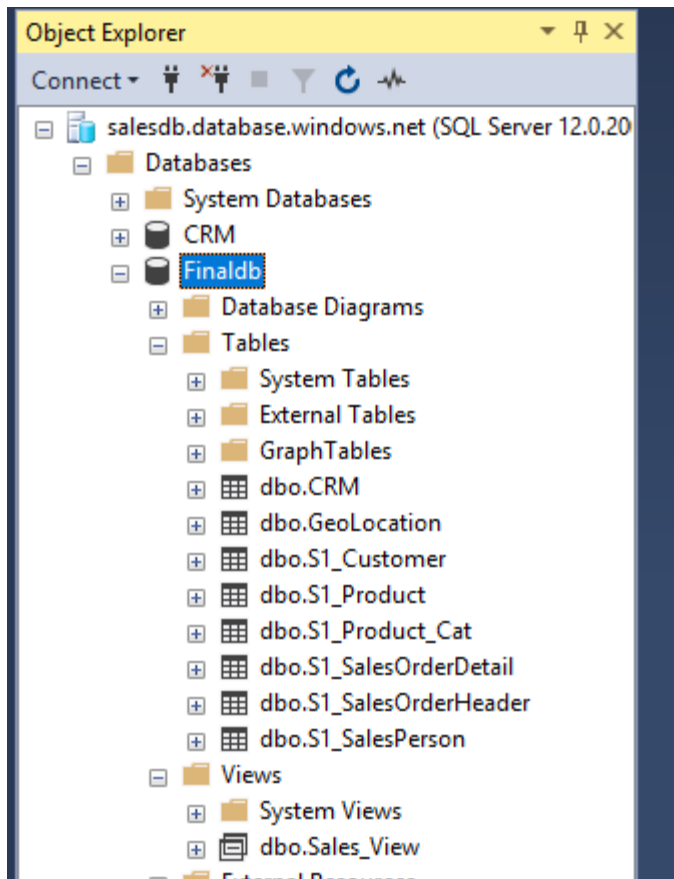
Εικόνα 4.29: Pipeline Completion



Εικόνα 4.30: Deployment Pipeline

4.5 Κατασκευή Τελικής Βάσης

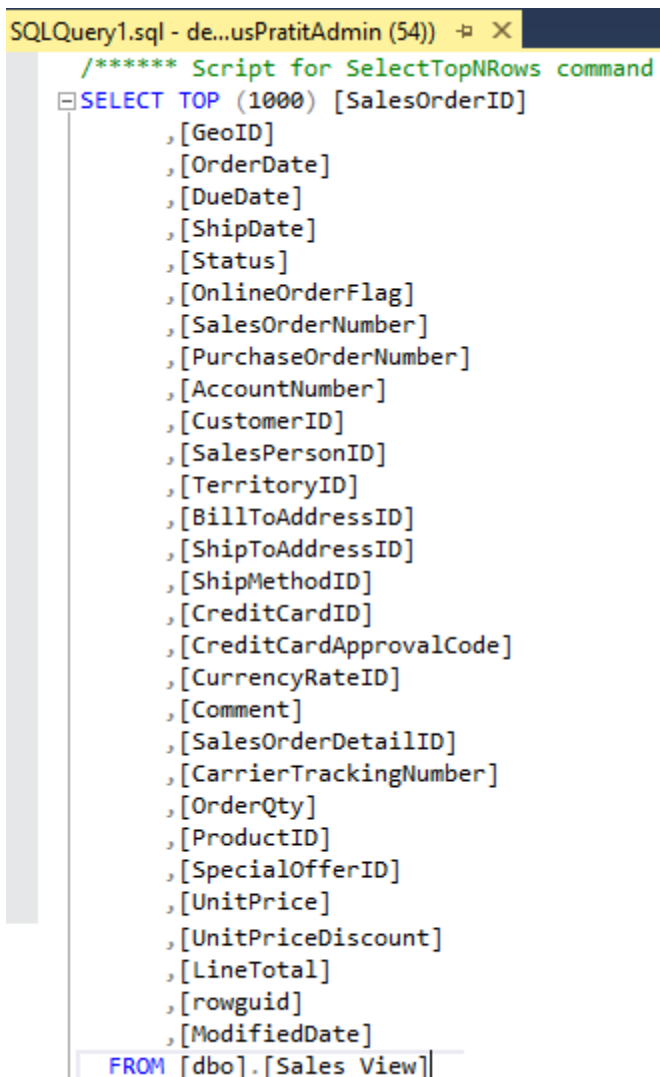
Μετά την ολοκλήρωση της ETL διαδικασίας στο Azure έχουμε την παρακάτω δομή στη βάση μας την οποία βλέπουμε μέσω του SSMS και αφού συνδεθούμε στην αντίστοιχη βάση δεδομένων.



Εικόνα 4.31: Τελική Δομή Βάσης

Γενικότερα στην ανάλυση, όταν έχουμε να διαχειριστούμε μεγάλο όγκο δεδομένων συνήθως αποφεύγεται η άμεση χρήση των βάσεων δεδομένων που περιέχουν δεδομένα σε λεπτομερειακή μορφή. Αυτό συμβαίνει καθώς περίπλοκα query που θα εκτελούνται στην βάση δεδομένων θα επιστρέφουν αποτελέσματα με μεγάλη καθυστέρηση κάτι που δεν βοηθάει στην ομαλή λειτουργία των διαδραστικών αναφορών. Επομένως τα δεδομένα που η εκάστοτε επιχείρηση αποφασίζει να χρησιμοποιήσει στα συστήματα BI πρέπει πρώτα να μετατραπούν σε μια μορφή συγκεντρωτικών δεδομένων. Για την μετατροπή αυτή υπάρχουν δύο επιλογές. Η μία είναι η κατασκευή κύβων δηλαδή ενός Data Warehouse που θα περιέχει πληροφορία γεγονότων για παράδειγμα πωλήσεις ενός συγκεκριμένου υποκαταστήματος για τον μήνα Δεκέμβρη. Είτε η κατασκευή ενός πίνακα με χρήση SQL ο οποίος θα φέρει πληροφορία από πολλαπλούς πίνακες. Στα πλαίσια της εργασίας αυτής χρησιμοποιήθηκε η δεύτερη μέθοδος. Δημιουργήθηκαν δηλαδή κάποιοι πίνακες (όψεις) οι οποίοι θα εμπεριέχουν πληροφορία η οποία είναι απαραίτητη για τα

διαγράμματα που θα κατασκευαστούν. Κατασκευάσαμε λοιπόν το Sales View στο οποίο συνδυάζεται πληροφορία από τους πίνακες πωλήσεων, τοποθεσίας, πελατών και μας δίνει πρόσβαση άμεσα σε πληροφορία σχετικά με παραγγελίες τις οποίες αργότερα θα απεικονίσουμε σε διαγράμματα. Για παράδειγμα μπορούμε να πάρουμε εύκολα πληροφορία για το πότε έγινε μια παραγγελία, πόσο κόστισε και ποιος ήταν ο παραλήπτης αυτής. Να επισημάνουμε ότι ο κώδικας SQL με τον οποίο κατασκευάστηκε το view αυτό περιλαμβάνεται στο παράρτημα Ι και γράφτηκε στο εργαλείο SSMS.



```
/****** Script for SelectTopNRows command
SELECT TOP (1000) [SalesOrderID]
, [GeoID]
, [OrderDate]
, [DueDate]
, [ShipDate]
, [Status]
, [OnlineOrderFlag]
, [SalesOrderNumber]
, [PurchaseOrderNumber]
, [AccountNumber]
, [CustomerID]
, [SalesPersonID]
, [TerritoryID]
, [BillToAddressID]
, [ShipToAddressID]
, [ShipMethodID]
, [CreditCardID]
, [CreditCardApprovalCode]
, [CurrencyRateID]
, [Comment]
, [SalesOrderDetailID]
, [CarrierTrackingNumber]
, [OrderQty]
, [ProductID]
, [SpecialOfferID]
, [UnitPrice]
, [UnitPriceDiscount]
, [LineTotal]
, [rowguid]
, [ModifiedDate]
FROM [dbo].[Sales_View]
```

Εικόνα 4.32: Όψη Sales View

Στον πίνακα της εικόνας 4.32 έχουμε φέρει πληροφορία από τον πίνακα πωλήσεων, τοποθεσίας, προϊόντων και πελατών ώστε να έχουμε σε ένα πίνακα την πληροφορία που θα εμφανίσουμε αργότερα σε αναφορές και να μην χρειάζεται να κάνει το σύστημα διαφορετικά

ερωτήματα (query) σε κάθε πίνακα ξεχωριστά. Για παράδειγμα αν θέλουμε να κατασκευάσουμε ένα διάγραμμα με τις πωλήσεις ανά τοποθεσία και πελάτη θα χρησιμοποιήσουμε πληροφορία που έχει μόνο ο συγκεκριμένος πίνακας κάνοντας έτσι την αναφορά μας πολύ πιο γρήγορη.

Στον πίνακα που έχουμε με τα CRM Δεδομένα θέλουμε και εκεί να κάνουμε κάποιους μετασχηματισμούς για να πάρουμε την πληροφορία που χρειαζόμαστε

Opp_BU	Opp_Status	Opp_Name	Opp_Salesman	Opp_Start_Date	Opp_End_Date	New_Amount	Client_ID	Opp_Id
BU7	Completed	Opportunity22	salesman9	4/11/2010 12:00:00 πμ	8/2/2011 12:00:00 πμ	56,0331335149864	22002	22
BU7	Completed	Opportunity23	salesman9	3/10/2011 12:00:00 πμ	17/1/2012 12:00:00 πμ	65,5226158038147	23001	23
BU7	Completed	Opportunity24	salesman9	14/3/2012 12:00:00 πμ	23/5/2012 12:00:00 πμ	42,0248501362398	23002	24
BU7	Completed	Opportunity25	salesman9	23/4/2014 12:00:00 πμ	6/7/2014 12:00:00 πμ	67,7820163487739	24001	25
BU7	Completed	Opportunity26	salesman9	30/3/2013 12:00:00 πμ	8/5/2013 12:00:00 πμ	72,3008174386921	24002	26
BU7	Completed	Opportunity27	salesman9	9/4/2013 12:00:00 πμ	22/6/2013 12:00:00 πμ	27,1128065395095	25001	27
BU7	Completed	Opportunity28	salesman9	27/5/2014 12:00:00 πμ	18/7/2014 12:00:00 πμ	14,0082833787466	25002	28
BU7	Completed	Opportunity29	salesman9	6/9/2013 12:00:00 πμ	3/1/2014 12:00:00 πμ	70,041416893733	26001	29
BU7	Completed	Opportunity30	salesman9	22/4/2011 12:00:00 πμ	8/8/2011 12:00:00 πμ	40,6692098092643	26002	30
BU7	Completed	Opportunity31	salesman9	1/8/2014 12:00:00 πμ	15/10/2014 12:00:00 πμ	56,0331335149864	27001	31

Εικόνα 4.33: Πίνακας CRM

Κρατάμε πληροφορία μόνο για το opportunity id, το start and end date και την ποσότητα. Αυτός ο πίνακας θα χρησιμευθεί σαν ενδιάμεσος πίνακας.

Opp_ID	Start_Date	End_Date	Quantity
1	30/5/2012 12:00:00 πμ	6/8/2012 12:00:00 πμ	120
2	23/11/2010 12:00:00 πμ	15/3/2011 12:00:00 πμ	1600
3	16/4/2013 12:00:00 πμ	23/5/2013 12:00:00 πμ	1200
4	26/11/2013 12:00:00 πμ	13/3/2014 12:00:00 πμ	313
5	7/9/2013 12:00:00 πμ	12/10/2013 12:00:00 πμ	800
6	3/5/2013 12:00:00 πμ	7/8/2013 12:00:00 πμ	600
7	2/1/2012 12:00:00 πμ	17/4/2012 12:00:00 πμ	3150
8	16/6/2014 12:00:00 πμ	6/10/2014 12:00:00 πμ	650
9	26/9/2014 12:00:00 πμ	22/12/2014 12:00:00 πμ	7,62295081967213
10	6/7/2014 12:00:00 πμ	15/10/2014 12:00:00 πμ	4,75409836065574

Εικόνα 4.34: CRM Data

Στη συνέχεια κατασκευάζουμε τον πίνακα CRM Data Break down ο οποίος περιέχει πληροφορία κατακερματισμένη σε ημέρες. Δηλαδή για κάθε ευκαιρία έχουμε την ημερήσια απόδοση που μας δίνει.

Opp_ID	Start_Date	End_date	Quantity	Daily_Qty	Ref_Date
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Τετάρτη, 30 Νοεμβρίου 2011
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Πέμπτη, 1 Δεκεμβρίου 2011
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Παρασκευή, 2 Δεκεμβρίου 2011
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Σάββατο, 3 Δεκεμβρίου 2011
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Κυριακή, 4 Δεκεμβρίου 2011
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Δευτέρα, 5 Δεκεμβρίου 2011
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Τρίτη, 6 Δεκεμβρίου 2011
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Τετάρτη, 7 Δεκεμβρίου 2011
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Πέμπτη, 8 Δεκεμβρίου 2011
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Παρασκευή, 9 Δεκεμβρίου 2011

Εικόνα 4.35: CRM Data breakdown

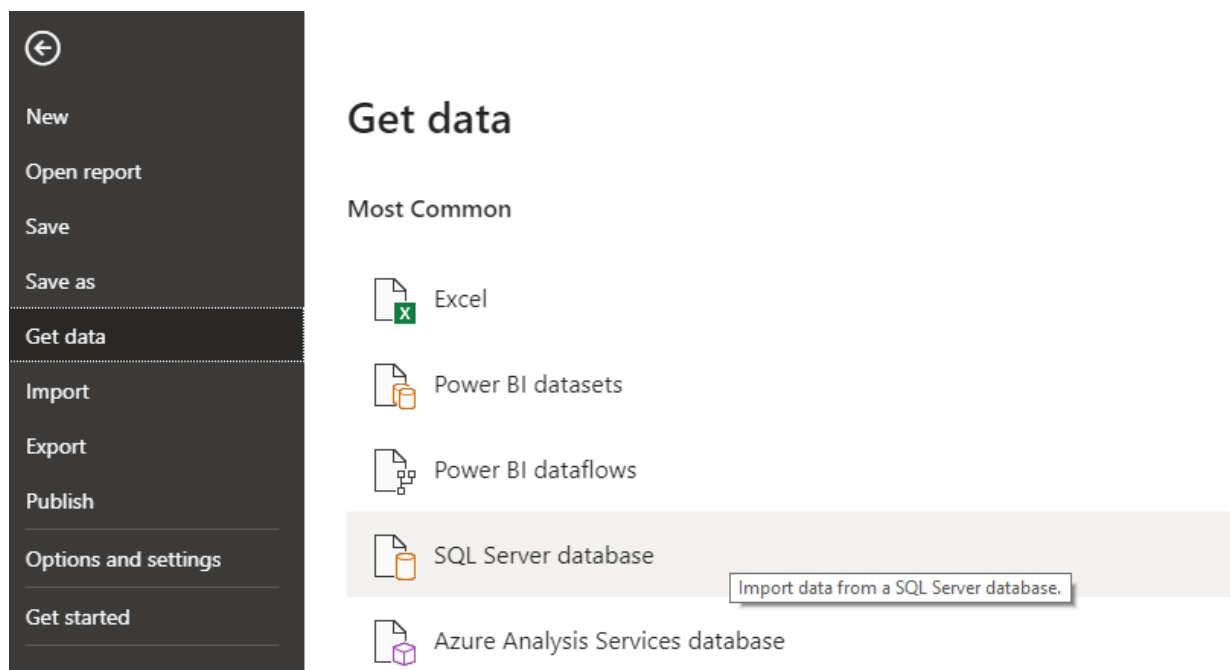
Τέλος κατασκευάζουμε την όψη που θα χρησιμοποιούμε για τις αναφορές και αφορά το κομμάτι του CRM.

Opp_ID	Start_Date	End_Date	Quantity	CRM_Value	Ref_Date	BU	BU_Status	BU_Name	BU_Salesman	Client_ID
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Πέμπτη, 1 Δεκεμβρίου 2011	BU9	Completed	Opportunity266	salesman3	25008
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Σάββατο, 1 Δεκεμβρίου 2012	BU9	Completed	Opportunity266	salesman3	25008
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Κυριακή, 1 Δεκεμβρίου 2013	BU9	Completed	Opportunity266	salesman3	25008
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Δευτέρα, 1 Δεκεμβρίου 2014	BU9	Completed	Opportunity266	salesman3	25008
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Τρίτη, 1 Δεκεμβρίου 2015	BU9	Completed	Opportunity266	salesman3	25008
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Πέμπτη, 1 Δεκεμβρίου 2016	BU9	Completed	Opportunity266	salesman3	25008
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Παρασκευή, 1 Δεκεμβρίου 2017	BU9	Completed	Opportunity266	salesman3	25008
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Σάββατο, 1 Δεκεμβρίου 2018	BU9	Completed	Opportunity266	salesman3	25008
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Κυριακή, 1 Δεκεμβρίου 2019	BU9	Completed	Opportunity266	salesman3	25008
266	Τετάρτη, 30 Νοεμβρίου 2011	Σάββατο, 28 Ιανουαρίου 2012		0,00173529411764706	Τρίτη, 1 Δεκεμβρίου 2020	BU9	Completed	Opportunity266	salesman3	25008

Εικόνα 4.36: Όψη CRM Final

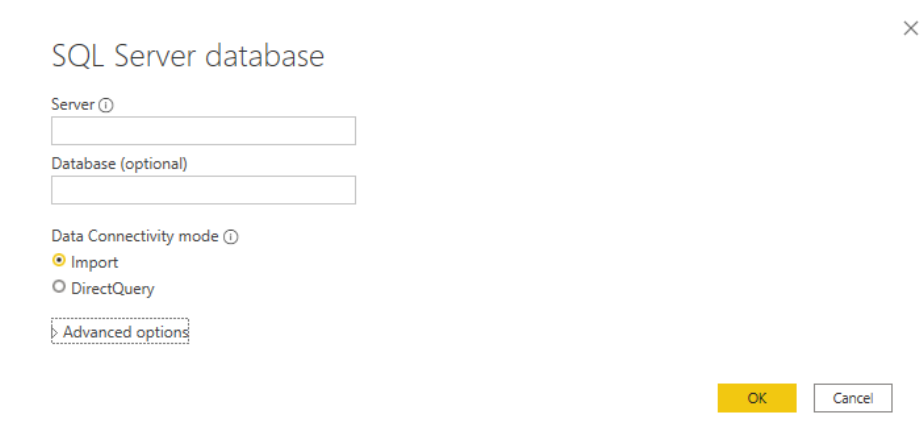
4.6 Import Δεδομένων στο Power BI

Όλη μας η ανάλυση θα γίνει στο Microsoft Power BI. Για να ξεκινήσουμε, αρχικά κατεβάζουμε το Power BI Desktop από το site της Microsoft <https://powerbi.microsoft.com/en-us/downloads/>. Αφού κάνουμε εγκατάσταση και ανοίξουμε το πρόγραμμα σειρά έχει να κάνουμε import της βάση μας στο πρόγραμμα για να ξεκινήσουμε την ανάλυση. Κατευθυνόμαστε στο File – Get Data – SQL Databases



Εικόνα 4.37: Import Data to Power BI

Στην οθόνη που εμφανίζεται βάζουμε το όνομα του Server της βάσης μας στο Azure το όνομα της Βάσης που θέλουμε να φορτώσουμε και πατάμε **OK**.



Εικόνα 4.38: Connection to Server

Επιλέγουμε τους πίνακες που θέλουμε να φορτώσουμε στο Power BI και στην συνέχεια επιλέγουμε **Transform Data**.

The screenshot shows the Power BI Navigator interface. On the left, the 'Display Options' pane shows a tree view with 'salesdb.database.windows.net [2]' expanded, then 'CRM [2]', and finally 'CRM\$' selected. Below this, 'Sales2020' is also visible. On the right, the 'CRM\$' table is previewed with the following data:

Opp BU	Opp Status	Opp Name	Opp Salesman	Opp Start Date
BU6	Completed	Opportunity1	salesman6	29/5/2015 12:0
BU6	Mature	Opportunity2	salesman6	21/4/2015 12:0
BU6	Completed	Opportunity3	salesman6	29/5/2015 12:0
BU6	Completed	Opportunity4	salesman6	29/1/2015 12:0
BU6	Completed	Opportunity5	salesman6	29/5/2015 12:0
BU6	Completed	Opportunity6	salesman2	29/5/2015 12:0
BU6	Completed	Opportunity7	salesman6	29/5/2015 12:0
BU6	Completed	Opportunity8	salesman6	29/5/2015 12:0
BU6	Completed	Opportunity9	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity10	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity11	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity12	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity13	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity14	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity15	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity16	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity17	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity18	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity19	salesman6	17/2/2015 12:0
BU6	Completed	Opportunity20	salesman6	17/2/2015 12:0
BU7	Completed	Opportunity21	salesman4	30/6/2015 12:0
BU7	Completed	Opportunity22	salesman9	9/5/2015 12:0
BU7	Completed	Opportunity23	salesman9	9/5/2015 12:0

At the bottom of the interface, there are three buttons: 'Select Related Tables', 'Load', and 'Transform Data' (which is highlighted with a red box), and a 'Cancel' button.

Εικόνα 4.39: Transform/Load Data Power BI

Σε αυτό το στάδιο μπορούμε να κάνουμε οποιοδήποτε μετασχηματισμό στα Data μας πριν τα φορτώσουμε στο πρόγραμμα, όπως για παράδειγμα να αλλάξουμε το όνομα μιας στήλης, να αφαιρέσουμε null τιμές ή ακόμα και να δημιουργήσουμε μια custom στήλη σε ένα πίνακα. Για την ώρα δεν θα πραγματοποιήσουμε κάποιο μετασχηματισμό γιατί θα γίνει παρακάτω, οπότε στην επόμενη οθόνη που θα εμφανιστεί θα πατήσουμε **Load**. Τα δεδομένα φορτώνονται στο πρόγραμμα δημιουργούνται οι συνδέσεις μεταξύ τους και το σχήμα της βάσης και είμαστε έτοιμη για να ξεκινήσουμε το κομμάτι της ανάλυσης.

4.7 Οπτικοποίηση και KIP's

Το πιο σημαντικό στάδιο σε ένα έργο Business Intelligence είναι αυτό της οπτικοποίησης. Σε αυτό το στάδιο γίνεται η απεικόνιση της πληροφορίας στον τελικό χρήστη με σκοπό την εξαγωγή συμπερασμάτων σχετικά με την πορεία της επιχείρησης. Το τι πληροφορία απεικονίζουμε, καθώς και ο τρόπος με τον οποίο αυτή παρουσιάζεται είναι αυτό που μπορεί να καθορίσει την επιτυχία ενός έργου BI. Πρέπει να έχει ακολουθηθεί σωστή ανάλυση της επιχείρησης καθώς και των απαιτήσεων που έχει από το έργο ώστε να προκύψουν οι κατάλληλοι δείκτες που θα θέλει να παρακολουθεί η επιχείρηση. Στόχος είναι άτομα σε θέσεις διοικητικές χωρίς την γνώση πληροφορικής να μπορούν να είναι σε θέση να κατανοήσουν την πληροφορία που θα εμφανίζεται στις αναφορές και να μπορούν εύκολα να εξάγουν συμπεράσματα μέσω αυτών. Τα συμπεράσματα αυτά είναι που θα οδηγήσουν με την σειρά τους στη λήψη στρατηγικών αποφάσεων για την επιχείρηση. Τα εργαλεία απεικόνισης της πληροφορίας, λοιπόν, ξεπερνούν τα στατικά γραφήματα που χρησιμοποιούνταν στα υπολογιστικά φύλλα του Excel έως τώρα, προβάλλοντας δεδομένα με πιο εξελιγμένους τρόπους, όπως είναι τα ραβδογράμματα, οι πίτες, τα ντόνατς, οι χάρτες θερμότητας (heat maps), τα ιστογράμματα και πολλά άλλα. Τα παραπάνω διαγράμματα μπορούν επίσης να έχουν διαδραστικότητα και αλληλεπίδραση με τα υπόλοιπα που περιλαμβάνει μια αναφορά, για παράδειγμα πατώντας σε ένα κομμάτι μιας πίτας τα υπόλοιπα γραφήματα θα ανανεώνονται και θα δείχνουν πληροφορίες οι οποίες σχετίζεται με το συγκεκριμένο στοιχείο που επιλέχθηκε.

Ένα ακόμα χαρακτηριστικό των σύγχρονων αυτών αναφορών είναι η δυνατότητα εφαρμογής φίλτρων όπως για παράδειγμα φίλτρων χρόνου, τα οποία όπως θα δούμε και στην συνέχεια εμφανίζουν στον χρήστη την πληροφορία που αντιστοιχεί στο χρονικό διάστημα που έχει επιλεγεί. Καθώς ο χρήστης θα επιλέγει μια συγκεκριμένη ημερομηνία στο φίλτρο θα γίνεται αυτόματη αλλαγή των διαγραμμάτων που έχουν σαν δεδομένο τον χρόνο και θα εμφανίζουν πληροφορίες για αυτό το διάστημα. Αν για παράδειγμα σε μια αναφορά έχουμε ένα διάγραμμα πωλήσεων ως προς τον χρόνο, ένα άλλος διάγραμμα με το κέρδος των πωλήσεων ως προς τον χρόνο και τέλος ένα πίνακα με όλα τα προϊόντα ανά κατηγορίες και εφαρμόζαμε ένα φίλτρο χρόνου θα άλλαζαν μόνο τα δύο πρώτα διαγράμματα καθώς στο τελευταίο δεν θα υπήρχε η συσχέτιση του χρόνου.

Η ανάλυση μας θα χωριστεί σε τρία κομμάτια. Στην ανάλυση των δεδομένων πωλήσεων όπου θα δούμε κάποιους βασικούς δείκτες και διαγράμματα στην συνέχεια στην ανάλυση των

CRM δεδομένων όπου θα κάνουμε μια αντίστοιχη ανάλυση και τέλος σε μία συγκεντρωτική όπου θα περιλαμβάνει δεδομένα και από τις δύο πηγές. Τέλος, θα δούμε κάποια σενάρια προγνωστικής αναλυτικής.

4.7.1 Sales Analytics

Σαν αποτέλεσμα της ανάλυσης απαιτήσεων που κάναμε για τα δεδομένα του ERP, κατασκευάσαμε την παρακάτω αναφορά η οποία το περιέχει κάποιους βασικούς δείκτης απόδοσης (KPIs) , οι οποίοι αποδεικνύουν πόσο αποτελεσματικά μια εταιρεία επιτυγχάνει τους επιχειρηματικούς στόχους της καθώς και διαγράμματα μέσω των οποίων μπορούμε να παρατηρήσουμε την εξάρτηση κάποιων dimensions.

A. KPIs πωλήσεων: Έσοδα, κέρδος, παραγγελίες και συνολικός αριθμός τεμαχίων που πωλήθηκαν.

B. Φίλτρα στο dimension του χρόνου: Ανά ημέρα, μήνα και έτος σύμφωνα με τα οποία εμφανίζονται data για το αντίστοιχο διάστημα.

Γ. Έσοδα και αριθμός τεμαχίων που πωλήθηκαν ως προς το dimension του χρόνου (Area Chart): Σύγκριση των εσόδων με των αριθμό τεμαχίων που πωλήθηκαν το αντίστοιχο χρονικό διάστημα, ώστε να έχει η επιχείρηση πλήρη εικόνα της αλληλεξάρτησης των δύο αυτών οντοτήτων.

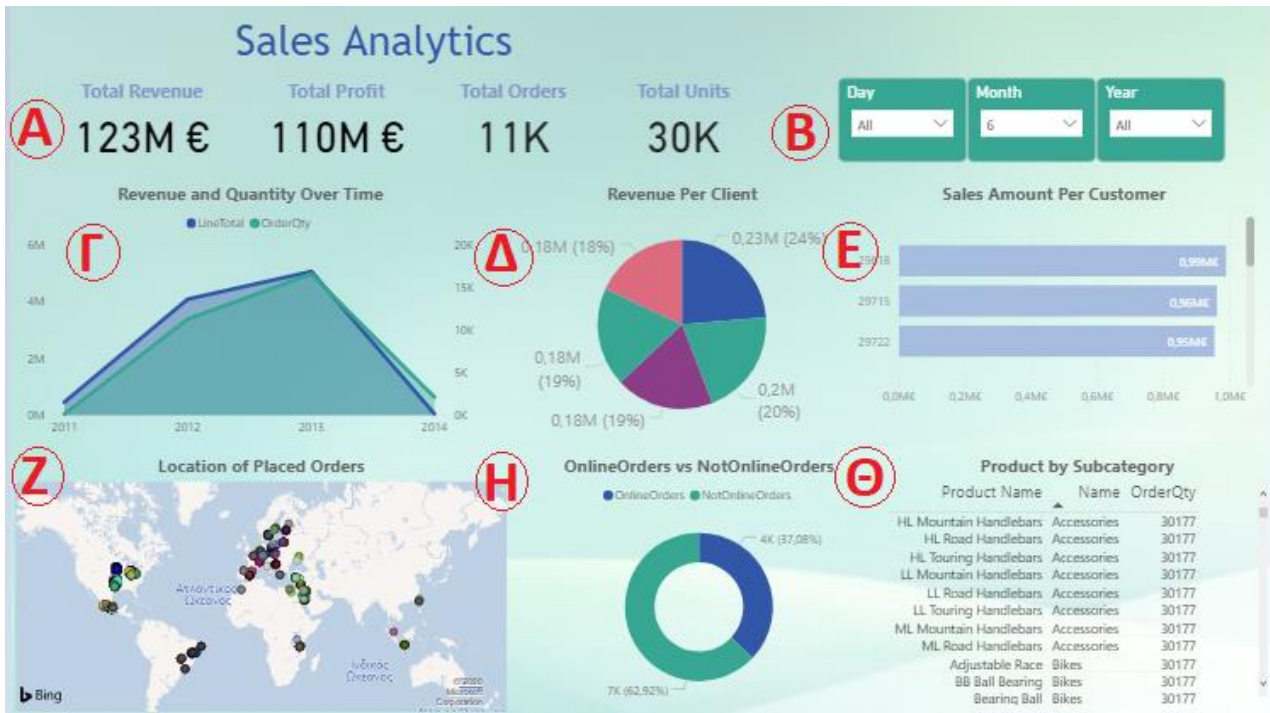
Δ. Ποσοστό κέρδους ανά πελάτη (Pie chart): Διάγραμμα που δείχνει πως κατανέμεται το κέρδος στους πελάτες.

Ε. Πωλήσεις ανά πελάτη (Stacked bar chart): Διάγραμμα που δείχνει τους top 7 πελάτες με βάση τα συνολικά έσοδα πωλήσεων.

Ζ. Τοποθεσία παραγγελιών (Map): Χάρτης που δείχνει τις περιοχές από τις οποίες πραγματοποιήθηκαν οι παραγγελίες.

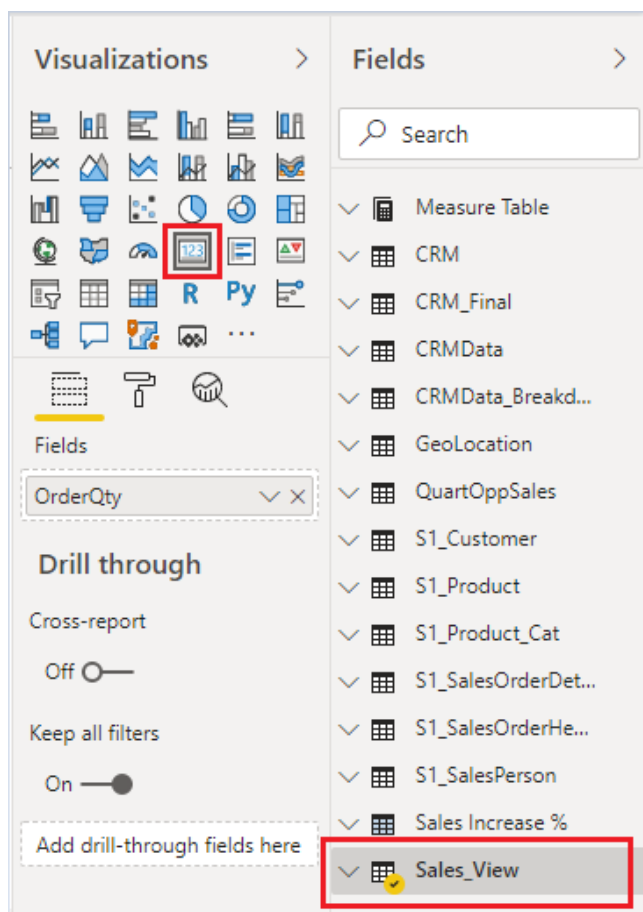
Η. Online vs On premise πωλήσεις (Donut chart): Με αυτόν τον τύπο διαγράμματος μπορούμε εύκολα να δούμε την προτίμηση των πελατών σε σχέση με τις online αγορές.

Θ. Προϊόντα ανά κατηγορία (Table): Διάγραμμα σε μορφή πίνακα στο οποίο βλέπουμε τα προϊόντα και την κατηγορία στην οποία ανήκουν.

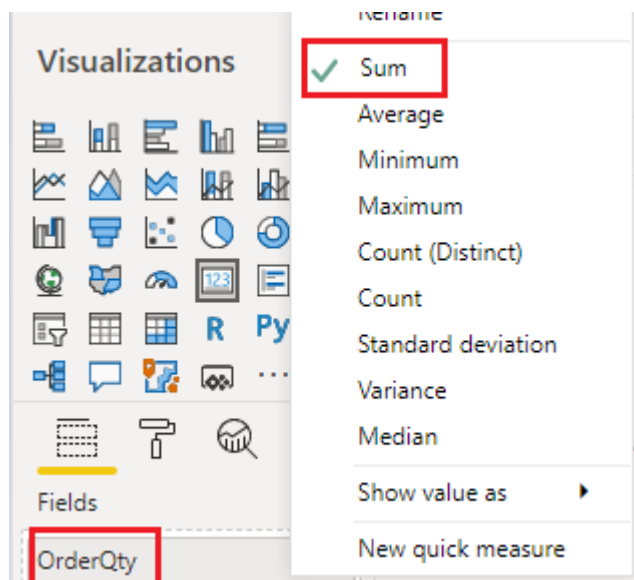


Εικόνα 4.40: Sales Report

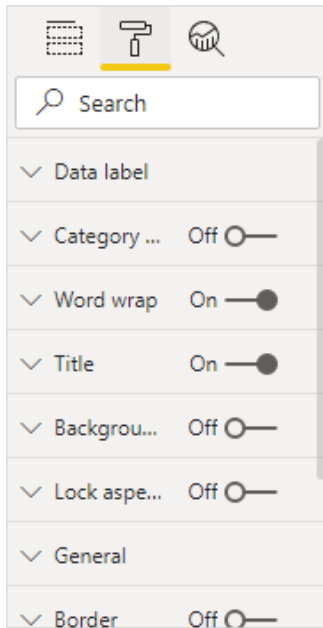
Στο σημείο αυτό θα παρουσιαστεί ο τρόπος κατασκευής της παραπάνω αναφοράς. Για την δημιουργία των KPIs ακολουθήσαμε την ίδια διαδικασία, οπότε θα δούμε ενδεικτικά το πως κατασκευάστηκε έναν από τους δείκτες. Για να κατασκευάσουμε το KPI των TotalUnits, αρχικά επιλέγουμε από τα διαθέσιμα visualizations αυτό της κάρτας (Card). Στο δεξί μέρος όπου έχουμε τους διαθέσιμους πίνακες βρίσκουμε το Sales View, πατάμε το βέλος για να δούμε τα columns που έχει ο πίνακας και επιλέγουμε το OrderQty. Η default επιλογή είναι να εφαρμόσει το Power BI αλγόριθμο Count. Στις εγγραφές που διαθέτει η στήλη OrderQty, θα βρούμε την συνολική τιμή, δηλαδή έναν αριθμό που αντιστοιχεί στον συνολικό αριθμό παραγγελιών της επιχείρησης. Σε περίπτωση που θα θέλαμε να εφαρμόσουμε κάποιον άλλο αλγόριθμο θα επιλέγαμε το βέλος δίπλα από το OrderQty, όπως βλέπουμε στην εικόνα 3-42, και στη συνέχεια θα επιλέγαμε τον αλγόριθμο που επιθυμούμε να εφαρμοστεί στις εγγραφές. Μετά από αυτές τις επιλογές η κάρτα μας θα έχει την τελική τιμή της και το μόνο που μένει είναι να αλλάξουμε την μορφοποίηση της, δηλαδή την γραμματοσειρά που χρησιμοποιεί, το περίγραμμα, τον τίτλο κα. (εικόνα 4-43).



Εικόνα 4.41: Κατασκευή KPI



Εικόνα 4.42: Αλλαγή αλγορίθμου



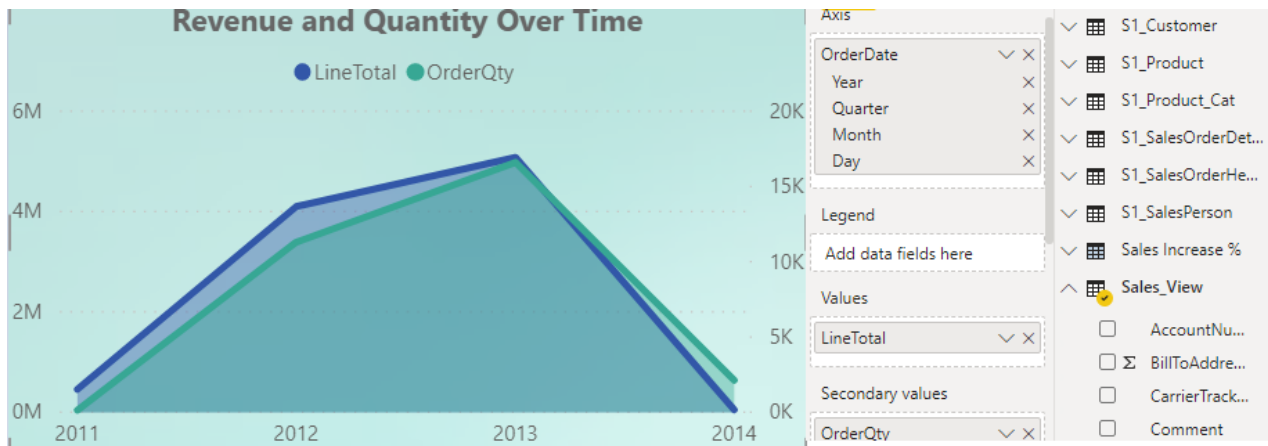
Εικόνα 4.43: Αλλαγή μορφοποίησης

Για την κατασκευή των φίλτρων αρχικά έπρεπε να εξάγουμε από την ημερομηνία παραγγελίας που έχουμε στο Sales View, τους μήνες τις μέρες και τα χρόνια. Για να πραγματοποιήσουμε αυτή την διαδικασία πηγαίνουμε στη σελίδα τροποποίησης των δεδομένων, επιλέγουμε την στήλη που θέλουμε, επιλέγουμε split columns και δημιουργούμε τρεις νέες στήλες οι οποίες θα είναι τα δεδομένα στα οποία θα βασιστούν τα φίλτρα μας. Επιλέγουμε το visualization slicer και επιλέγουμε την αντίστοιχη στήλη για το slicer του μήνα, της ημέρας και του έτους αντίστοιχα.

OrderQty	ProductID	SpecialOfferID	UnitPrice	UnitPriceDiscount	LineTotal	rowguid	ModifiedDate	OrderDate - Copy.1	Order_Month	Order_Year
3	716	1	28,8404 €	0 €	86,5212	7B08C20B-9A3F-4CFC-870B-E13375286740	31/5/2011 12:00:00 πμ	31	5	2011
1	771	1	2,039,994 €	0 €	2039,994	04C4DE91-5815-45D6-8670-F462719F8CE3	31/5/2011 12:00:00 πμ	31	5	2011
1	772	1	2,039,994 €	0 €	2039,994	5A74C7D2-E641-438E-A7AC-37BF23280301	31/5/2011 12:00:00 πμ	31	5	2011
2	773	1	2,039,994 €	0 €	4079,988	CE472532-A4C0-45BA-816E-EEFD3FD848B3	31/5/2011 12:00:00 πμ	31	5	2011
1	774	1	2,039,994 €	0 €	2039,994	80667840-F962-4EE3-96E0-AECA108E0D4F	31/5/2011 12:00:00 πμ	31	5	2011
3	714	1	28,8404 €	0 €	86,5212	E9D54907-E7B7-4969-80D9-76BA69F8A836	31/5/2011 12:00:00 πμ	31	5	2011
1	716	1	28,8404 €	0 €	28,8404	AA542690-BDCD-4CE5-89A0-C1BF82747725	31/5/2011 12:00:00 πμ	31	5	2011
6	709	1	5,7 €	0 €	34,2	AC769034-3C2F-495C-ASA7-3B71CDB25D4E	31/5/2011 12:00:00 πμ	31	5	2011
2	712	1	5,1865 €	0 €	10,373	06A66921-689F-4199-A912-DDAFD383472B	31/5/2011 12:00:00 πμ	31	5	2011
4	711	1	20,1865 €	0 €	80,746	0E371EE3-253E-48B0-8B13-83CF4224F972	31/5/2011 12:00:00 πμ	31	5	2011

Εικόνα 4.44: Split Columns

Για την κατασκευή του συγκριτικού διαγράμματος (γ εικόνα 3-40) όπως προείπαμε έχουμε επιλέξει το γραφικό Area Chart. Στον άξονα των υποποθετούμε την ημερομηνία παραγγελίας την οποία το BI την αναγνωρίζει σαν μια οντότητα η οποία αποτελείται από years, quarters, months και days. Στον άξονα των υποποθετούμε το Sum των εσόδων και το Sum των παραγγελιών για να δούμε πως σχετίζονται οι δύο αυτές έννοιες στο διάστημα του χρόνου.



Εικόνα 4.45: Διάγραμμα Revenue and Quantity

4.7.2 CRM Analytics

Σαν αποτέλεσμα της ανάλυσης απαιτήσεων που κάναμε για τα δεδομένα του CRM κατασκευάσαμε την παρακάτω αναφορά (εικόνα 4-46), η οποία μας δίνει πληροφορίες σε σχέση με τις ευκαιρίες (opportunities) για πωλήσεις που είχε η εταιρεία μας και την πορεία αυτών. Η αναφορά αυτή περιλαμβάνει τα εξής:

A. KPIs:

- Win Rate: Το ποσοστό των ευκαιριών που εξελίχθηκαν σε πραγματική πώληση.
- Loss Rate: Το ποσοστό των ευκαιριών που η εταιρεία έχασε.
- Won Revenue: Από τις ευκαιρίες που εξελίχθηκαν σε πωλήσεις υπολογίζουμε το συνολικό κέρδος.
- Lost Revenue: Από το σύνολο των ευκαιριών υπολογίζουμε το συνολικό κέρδος το οποίο η εταιρεία δεν κατάφερε να αποκτήσει.

B. Κέρδος που αποκτήθηκε σε σχέση με το εκτιμώμενο (Gauge): Διάγραμμα στο οποίο βλέπουμε το ποσό που κερδήθηκε σε σχέση με το εκτιμώμενο.

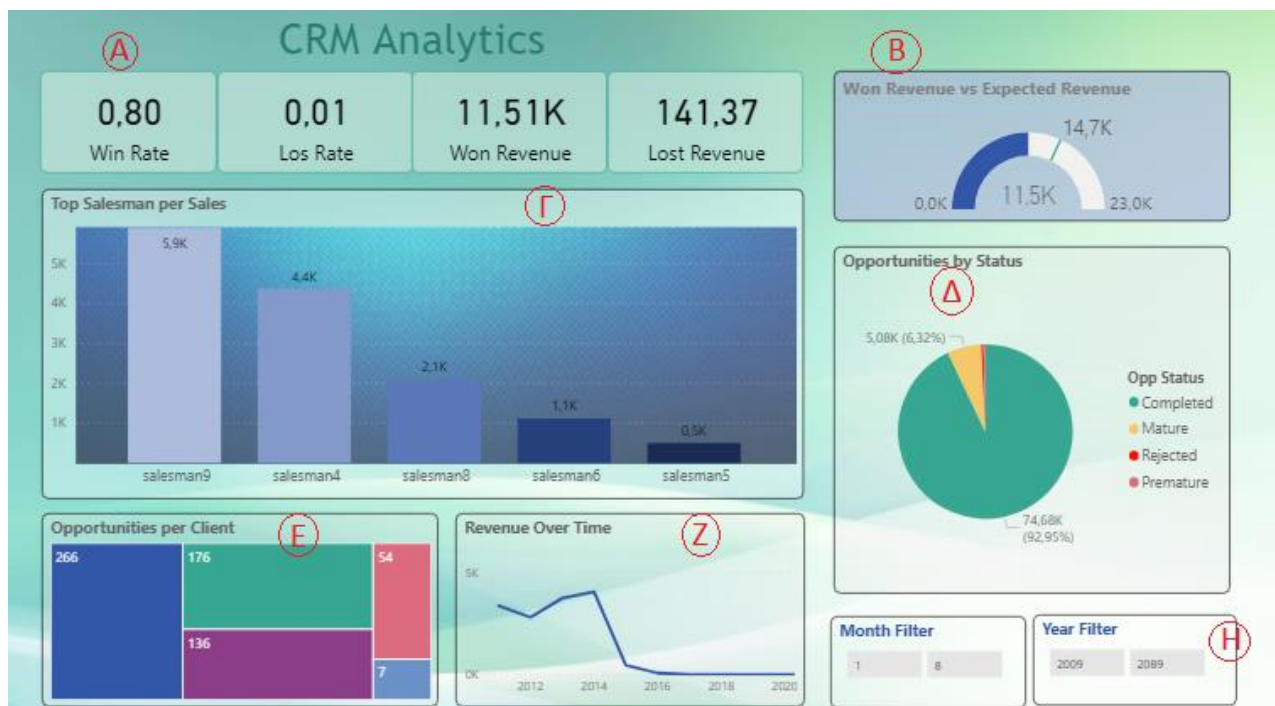
Γ. Καλύτεροι πωλητές με βάση το κέρδος (Stacked Column Chart): Διάγραμμα που δείχνει τους πέντε καλύτερους πωλητές με βάση το κέρδος που αποκτήθηκε.

Δ. Ευκαιρίες ανά κατηγορία (Pie chart): Διάγραμμα που δείχνει το ποσοστό που καλύπτει κάθε είδος ευκαιρίας.

Ε. Ευκαιρίες ανά πελάτη (Tree Map): Διάγραμμα που δείχνει τους top5 πελάτες, δηλαδή τους πελάτες που μας έδωσαν τις περισσότερες προσφορές για πώληση.

Ζ. Κέρδος ως προς τον χρόνο (Line Chart): Διάγραμμα που δείχνει στην πορεία του κέρδους σε βάθος χρόνου.

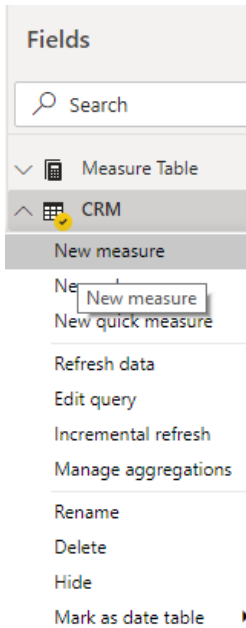
Η. Φίλτρα έτους και μήνα: Φίλτρα με τα οποία μπορούμε να εξάγουμε πληροφορία από την αναφορά σχετικά με την επιλογή μας.



Εικόνα 4.46: CRM Analytics Report

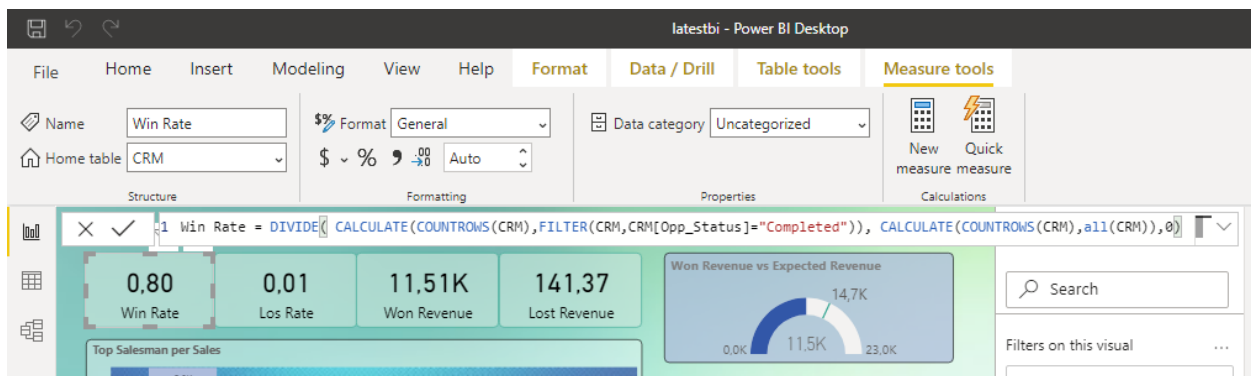
Για την κατασκευή του Win Rate δημιουργήσαμε ένα measurement μέσω του Power BI. Κάνοντας χρήση της γλώσσας DAX του Power BI κάνουμε δηλαδή έναν υπολογισμό (query) στα δεδομένα μας με τον εξής τρόπο:

1. Κάνουμε δεξί κλικ στον Πίνακα στον οποίο θέλουμε να δημιουργήσουμε το measurement και επιλέγουμε νέο measurement. Αυτομάτως μας ανοίγει ένα λευκό παράθυρο (editor) πάνω από την αναφορά μας στο οποίο μπορούμε να γράψουμε τον κώδικα σε γλώσσα Dax και να το δημιουργήσουμε (εικόνα 4-47)



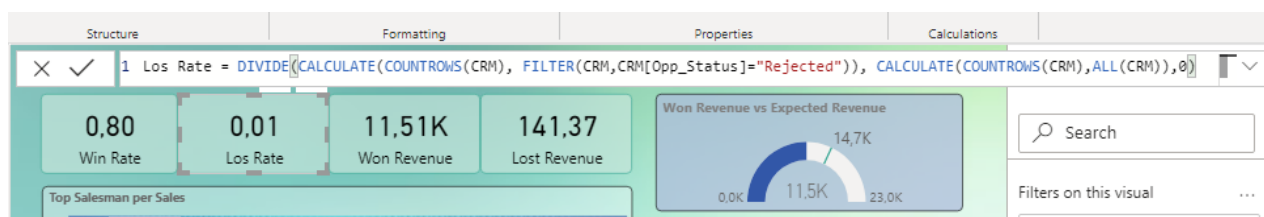
Εικόνα 4.47: New measurement

Ουσιαστικά, για να βρούμε το ποσοστό επιτυχίας πρέπει να υπολογίσουμε τις ευκαιρίες που μετατράπηκαν σε πωλήσεις. Άρα πρέπει να υπολογιστούν οι ευκαιρίες που έχουν status completed προς τις συνολικές ευκαιρίες * 100 για να βρούμε το ποσοστό επιτυχίας. Αντίστοιχη εργασία κάνουμε και για την δημιουργία του loss rate με την διαφορά ότι θα αθροίσουμε τις ευκαιρίες με status failed αντί για completed. Σε αυτό το σημείο να αναφέρουμε ότι όταν δημιουργούμε ένα νέο measurement δηλαδή, ένα προκατασκευασμένο άθροισμα μεγεθών ως προς κάποιο άλλο όπως είναι ο χρόνος ή μια περιοχή το Power BI ουσιαστικά κατασκευάζει εκείνη την στιγμή ένα κύβο κάνοντας τις συνδέσεις μεταξύ των πινάκων αυτόματα και τον χρησιμοποιεί κάθε φορά που χρησιμοποιούμε το measurement αυτό αντί να κάνει τον υπολογισμό από την αρχή καταναλώνοντας υπολογιστική ισχύ από το σύστημα. Αυτοματοποιεί επομένως την διαδικασία κατασκευής κύβων δίνοντας της δυνατότητα στον χρήστη να παράξει οποιαδήποτε στιγμή της



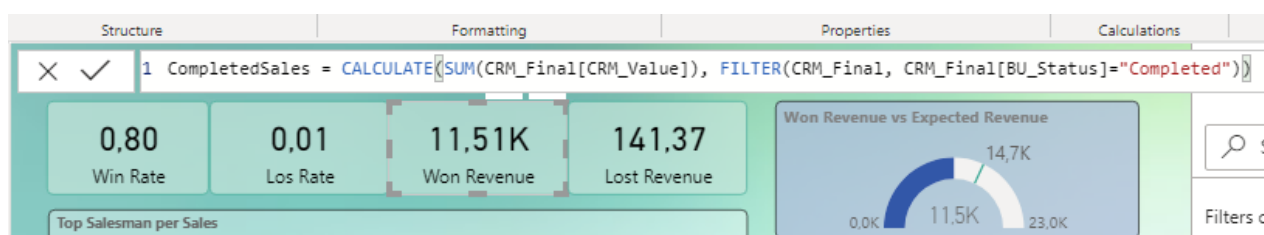
Εικόνα 4.48: Dax Win Rate

ανάλυσης συνδέσεις μεταξύ των δεδομένων του μη περιορίζοντάς τον στην κατασκευή ενός κύβου εξαρχής.

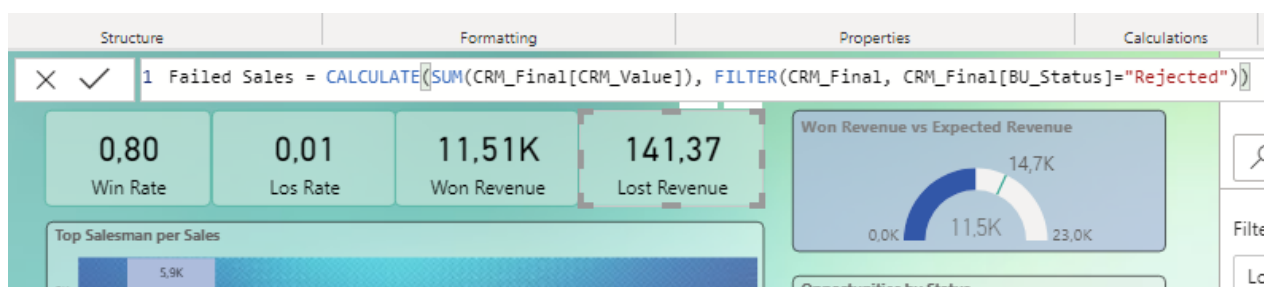


Εικόνα 4.49: Dax Loss Rate

Για την κατασκευή το Won Revenue και Lost Revenue αντίστοιχα αθροίσαμε το κέρδος κάθε ευκαιρίας με status completed και αντίστοιχα κάθε ευκαιρίας με status failed.

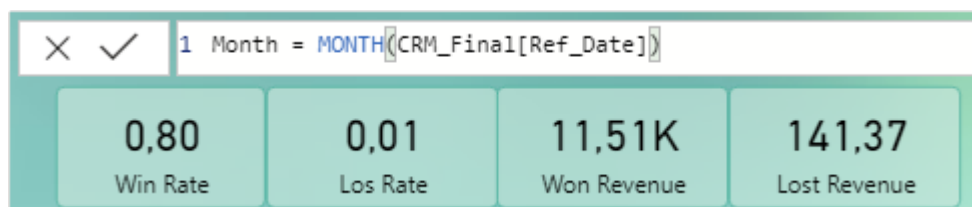


Εικόνα 4.50: Won Revenue



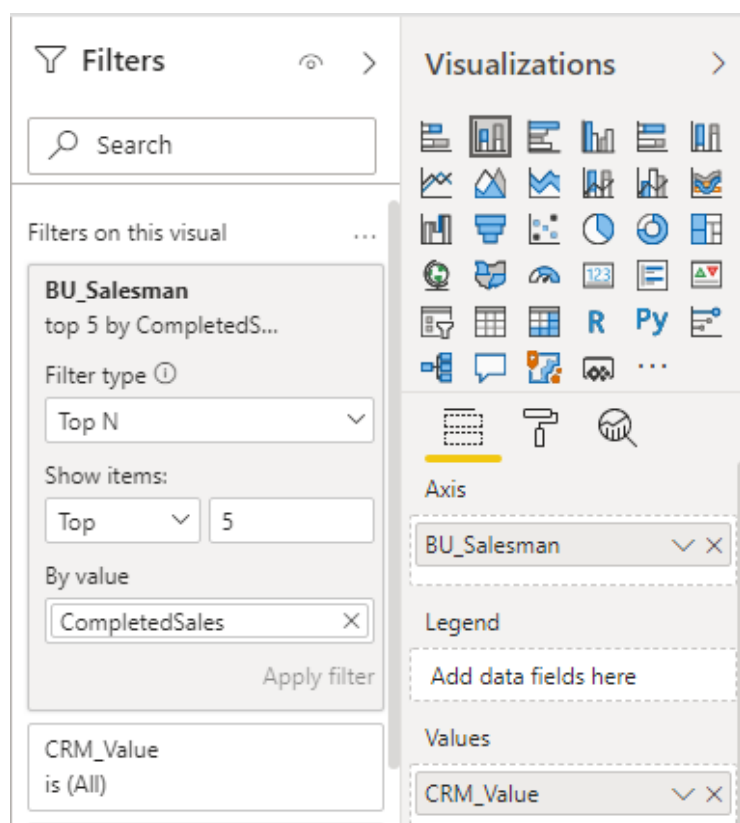
Εικόνα 4.51: Lost Revenue

Ένας άλλος τρόπος από αυτόν που δείξαμε για να κατασκευάσουμε την οντότητα του μήνα από μια ημερομηνία είναι και η κατασκευή ενός measurement. Γράφοντας κώδικα σε DAX μπορούμε πολύ απλά να εξάγουμε τον μήνα ή τον χρόνο από μια ημερομηνία χρησιμοποιώντας την.



Εικόνα 4.52: Month Measurement

Όπως έχουμε αναφέρει μέχρι στιγμής όταν κατασκευάζουμε ένα νέο διάγραμμα και κάνουμε για παράδειγμα ένα ερώτημα (query) για τους πωλητές με βάση τις πωλήσεις, το διάγραμμα θα γεμίσει με τις εγγραφές όλων των πωλητών που έχει ο πίνακός μας. Έτσι όμως η πληροφορία δεν θα είναι ευδιάκριτη για τον τελικό χρήστη. Πρέπει, λοιπόν με κάποιον τρόπο να μειώσουμε την πληροφορία που θα απεικονιστεί. Έστω λοιπόν ότι θέλουμε να κρατήσουμε τους Top 5 πωλητές με βάση τις πωλήσεις που πραγματοποίησαν. Για να το κάνουμε αυτό θα πρέπει να δημιουργήσουμε ένα νέο φίλτρο. Αυτή η διαδικασία όμως διαφέρει από αυτήν της κατασκευής φίλτρων για όλο το report καθώς αυτό το φίλτρο θα ανήκει στο συγκεκριμένο διάγραμμα μόνο. Η κατασκευή του γίνεται από το tab των Filters, στο οποίο θα εφαρμόσουμε Top 5 στα αποτελέσματα του ερωτήματος.



Εικόνα 4.53: Filter Top 5

4.8 Predictive Modeling and Advanced Analytics

Σε αυτό το σημείο συνδυάζουμε τις πληροφορίες που έχουμε και από τις δύο πηγές μας με στόχο τον εντοπισμό τάσεων, την πρόγνωση μελλοντικών πωλήσεων καθώς και την ανάλυση εκτιμώμενων στόχων όσον αναφορά ευκαιρίες πελατών σε σχέση με τις πραγματικές πωλήσεις και το κέρδος που αυτές αποφέρανε στην επιχείρησή μας. Γίνεται λοιπόν χρήση προγνωστικής αναλυτικής και ανάλυση των δεδομένων με σκοπό την εξαγωγή πιο σύνθετων συμπερασμάτων που θα αποφέρουν περισσότερες απαντήσεις στην επιχείρηση για το πώς κινείται στην αγορά.

A. KPIs πραγματικών αποτελεσμάτων σε σχέση με εκτιμώμενα:

- Πωλήσεις σε σχέση με τον στόχο (KPI): Με αυτό τον τύπο διαγράμματος μπορούμε να ορίσουμε ποια είναι η τιμή της μεταβλητής και ποια ήταν η εκτιμώμενη, έτσι ώστε να μας δείξει αν απέχουμε από τον στόχο μας ή να τον έχουμε ξεπεράσει. Στη συγκεκριμένη περίπτωση βλέπουμε ότι απέχουμε από τον στόχο μας κατά 18.69%.
- Ποσότητα πωληθέντων σε σχέση με τον στόχο (KPI): Στη συγκεκριμένη περίπτωση βλέπουμε ότι έχουμε ξεπεράσει τον στόχο μας για την ποσότητα τεμαχίων που θέλαμε να πουλήσουμε και μάλιστα ότι έχουμε ξεπεράσει τον στόχο μας κατά 70.61%.

B. Ολοκληρωμένες πωλήσεις σε σχέση με το σύνολο ευκαιριών (Gauge): Στο διάγραμμα αυτό βλέπουμε πόσες ήταν οι ολοκληρωμένες πωλήσεις σε σχέση με τις συνολικές ευκαιρίες.

Γ. Ποσοστό επιτυχίας επίτευξης πώλησης (Card): Ποσοστό που δείχνει την πιθανότητα μια ευκαιρία να γίνει πώληση.

Δ. Pipeline πωλήσεων ανά Quarter: Ανάλυση που μας δίνει το κέρδος που αποκτήσαμε από τις ευκαιρίες που έγιναν πωλήσεις σε σχέση με το εκτιμώμενο κέρδος που είχε καθοριστεί για την ευκαιρία.

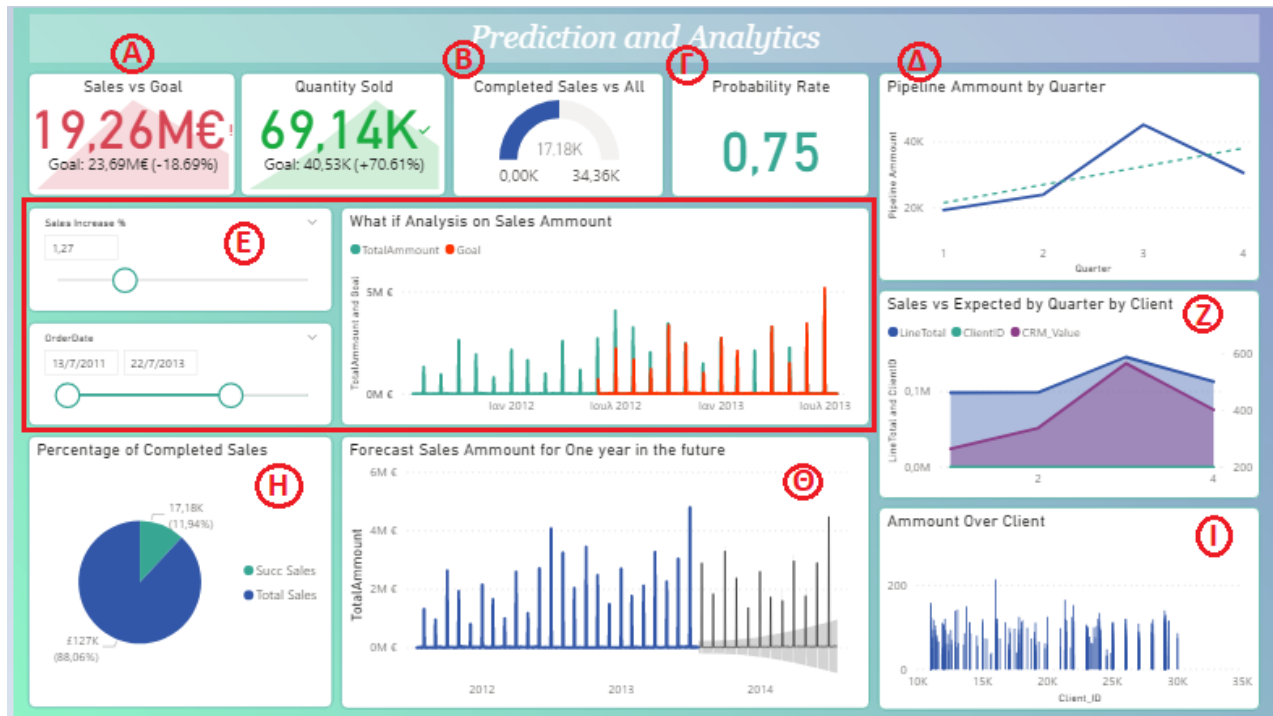
E. What if Analysis: Σενάριο προγνωστικής αναλυτικής σε σχέση με τις πωλήσεις που επιθυμούμε να επιτύχουμε.

Z. Πωλήσεις σε σχέση με εκτιμώμενες ανά Quarter και Πελάτη: Διάγραμμα που δείχνει τις πωλήσεις που πραγματοποιήθηκαν ανά πελάτη και τρίμηνο σε σχέση με τις εκτιμώμενες.

H. Ποσοστό πραγματοποιηθέντων πωλήσεων (pie chart): Πίτα που δείχνει πόσες πωλήσεις κάναμε σε σχέση με το σύνολο των ευκαιριών που είχε η εταιρεία.

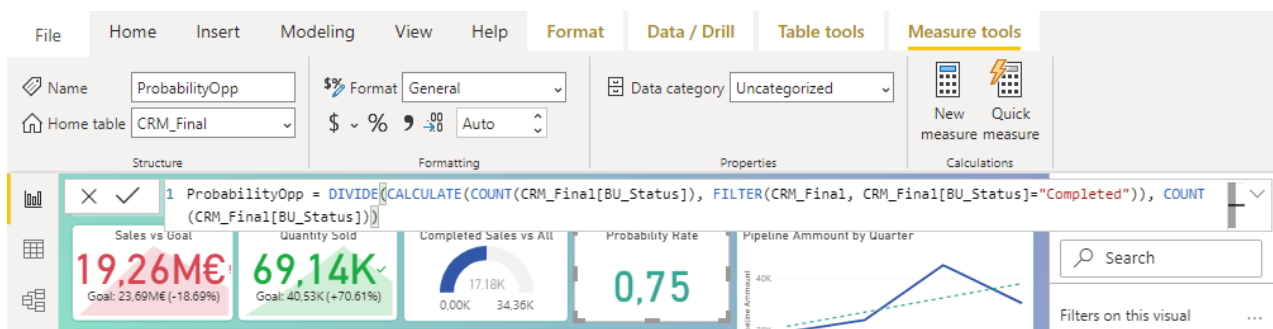
Θ. Forecast Πωλήσεων για τον επόμενο Χρόνο: Προγνωστική αναλυτική για την πρόβλεψη πωλήσεων του επόμενου χρόνου.

Ι. Πωλήσεις ανά Πελάτη: Διάγραμμα που δείχνει το πως κινήθηκαν οι πωλήσεις ανά πελάτη.



Εικόνα 4.54: Prediction and Analytics Report

Για την κατασκευή του Probability Rate φτιάχνουμε ένα νέο measurement και χρησιμοποιώντας την γλώσσα Dax παίρνουμε το άθροισμα των ευκαιριών που έγιναν completed προς το άθροισμα όλων των ευκαιριών πολλαπλασιαζόμενο με το 100.



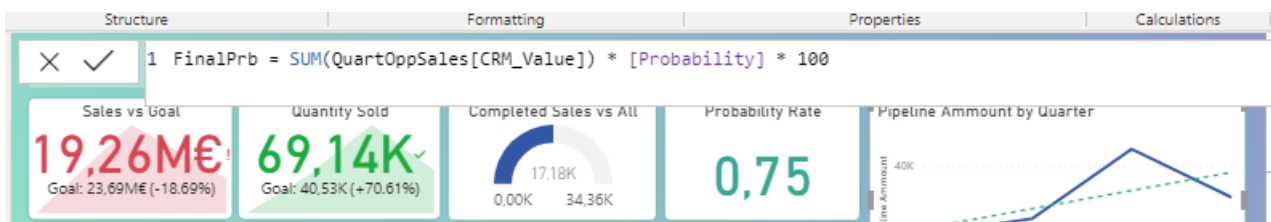
Εικόνα 4.55: Probability

Κατασκευή Pipeline Amount

Στόχος αυτής της ανάλυσης ήταν ο τελικός χρήστης να κατανοήσει την συσχέτιση μεταξύ των δεδομένων πωλήσεων και των εταιρικών σχέσεων και άρα να μπορέσει να εξάγει χρήσιμα συμπεράσματα από αυτά.

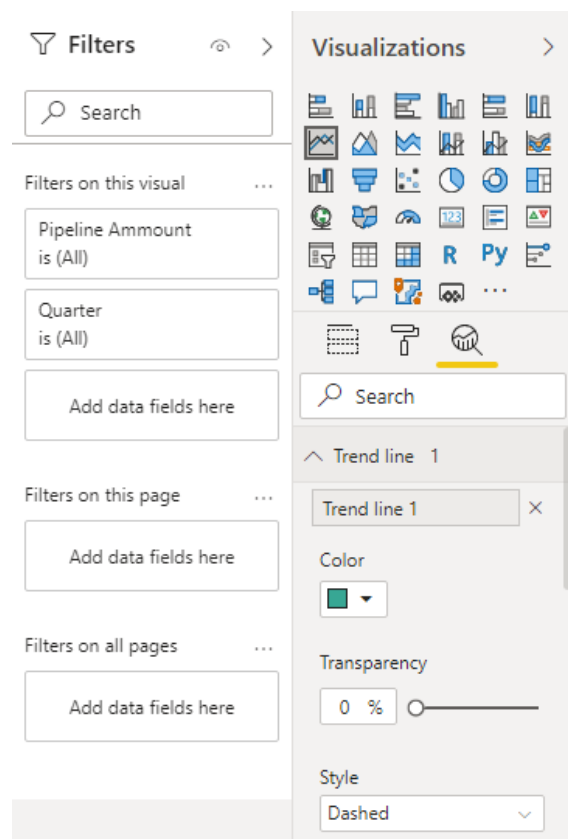
Για την κατασκευή αυτή αρχικά έπρεπε να φτιάξουμε ένα νέο πίνακα (κώδικας για τον οποίο υπάρχει στο παράρτημα Ι). Σε αυτό τον πίνακα θέλαμε να φέρουμε πληροφορία των crm και salesδεδομένων χωρισμένη ανά τρίμηνα για να μπορέσουμε να προσχωρήσουμε σε περαιτέρω αναλύσεις. Το pipeline amount ουσιαστικά θα εμφανίζει τα κέρδη των sales επί το probability rate που έχουμε ήδη κατασκευάσει και είναι μορφή ανάλυσης που γίνεται συχνά σε τέτοια είδους δεδομένα

Για την κατασκευή αυτού του amount θα χρησιμοποιήσουμε το ποσό των sales που έγιναν για κάθε τρίμηνο επί το αντίστοιχο probability rate εκείνης της περιόδου. Θα κατασκευάσουμε λοιπόν ένα νέο measurement το οποίο θα κάνει χρήση και του ήδη υπάρχοντος.

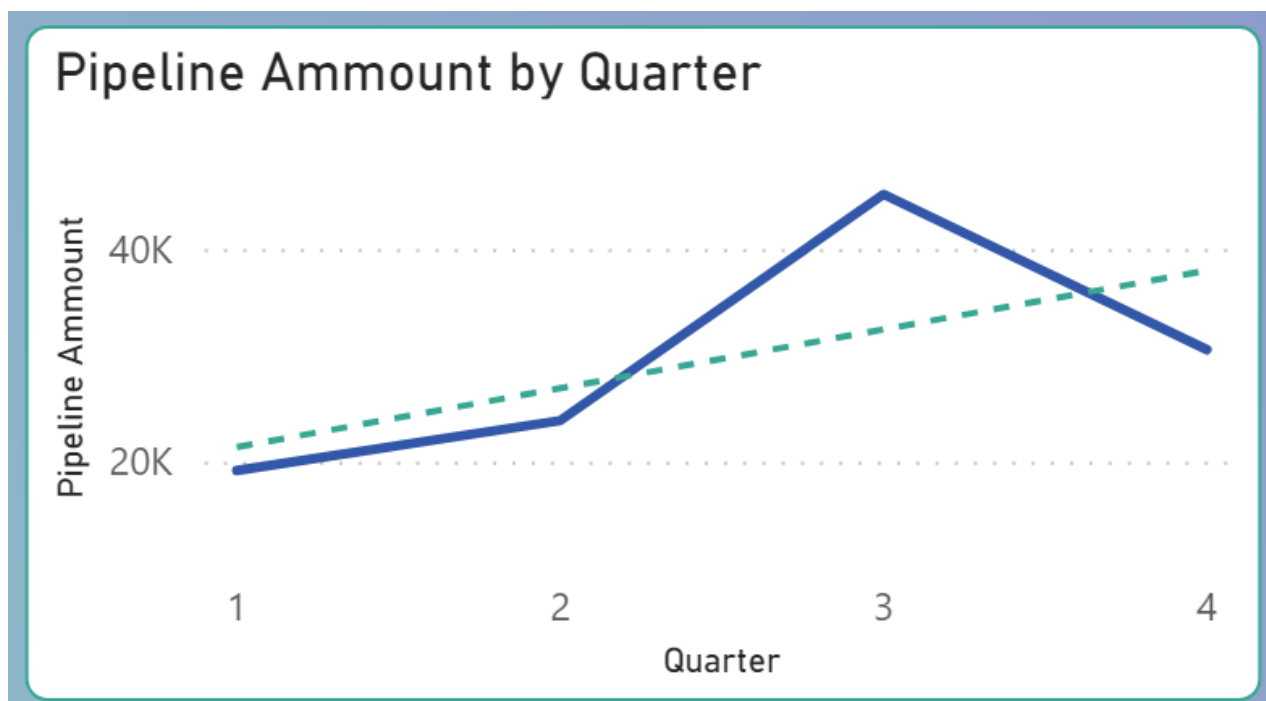


Εικόνα 4.56: Κατασκευή measurement για pipeline

Αφού κατασκευάσαμε το διάγραμμα μας, θέλουμε να προσθέσουμε ένα Trend Line για να δούμε πως κινούνται τα sales στο διάστημα του χρόνου. Για την κατασκευή του πηγαίνουμε στο tab των analytics και επιλέγουμε από τα build in μοντέλα ανάλυσης το Trend Line. (όπως βλέπουμε και στην εικόνα 4-57).



Εικόνα 4.57: Trend Line Creation

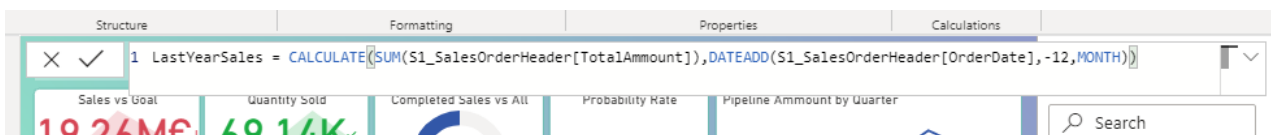


Εικόνα 4.58: Trend Line

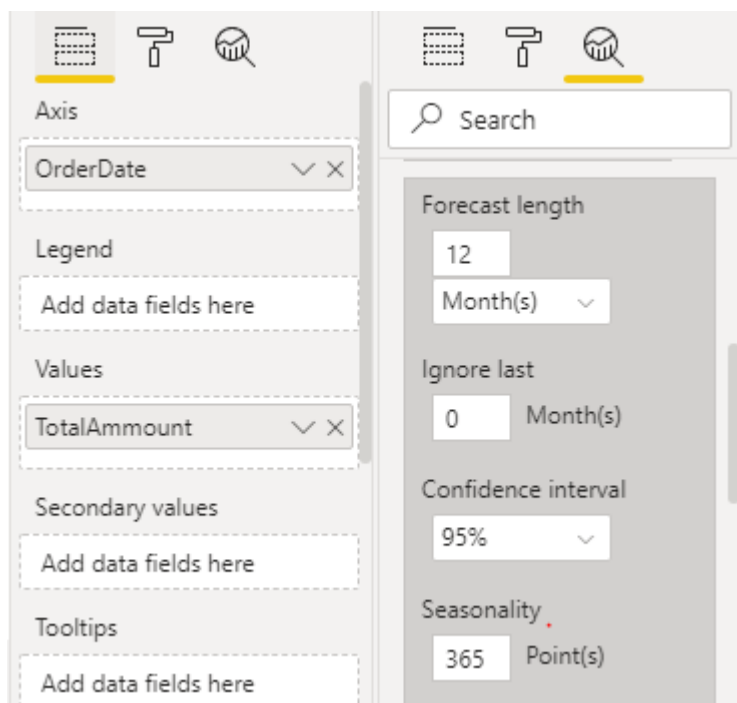
Βλέπουμε ότι η τάση μας παρουσιάζει μια ανοδική πορεία καθώς κινούμαστε στον άξονα του χρόνου.

Κατασκευή Forecasting

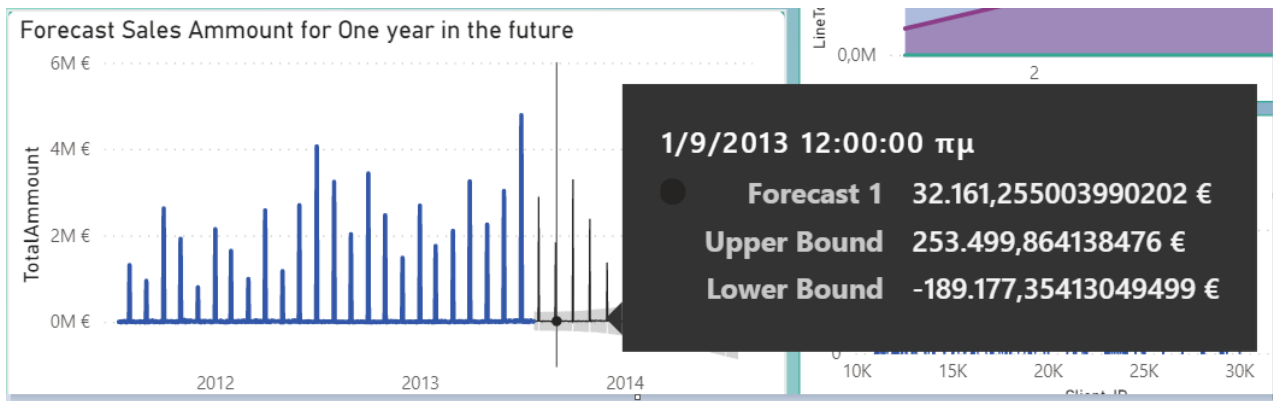
Είναι μια τεχνική ανάλυσης με την οποία μπορούμε να χρησιμοποιήσουμε δεδομένα του παρελθόντος για να κάνουμε μια πρόβλεψη μελλοντικών γεγονότων. Στην δική μας υλοποίηση θέλουμε να κάνουμε πρόβλεψη των πωλήσεων για τον επόμενο χρόνο με βάση τις πωλήσεις της προηγούμενης χρονιάς. Για αυτό το λόγο θα φτιάξουμε αρχικά ένα measurement για τις πωλήσεις της προηγούμενης χρονιάς δηλαδή για το έτος 2013 και θα το χρησιμοποιήσουμε το μοντέλο forecasting του Power BI για να βρούμε τις πωλήσεις που θα γίνουν κατά το έτος 2014. Επιλέγουμε forecasting για 12 μήνες μπροστά και τραβάμε καθημερινά δεδομένα (365 points) από τον πίνακα των πωλήσεων για να έχουμε μεγαλύτερη ακρίβεια στα αποτελέσματά μας.



Εικόνα 4.59: Πωλήσεις 2013



Εικόνα 4.60: Forecasting

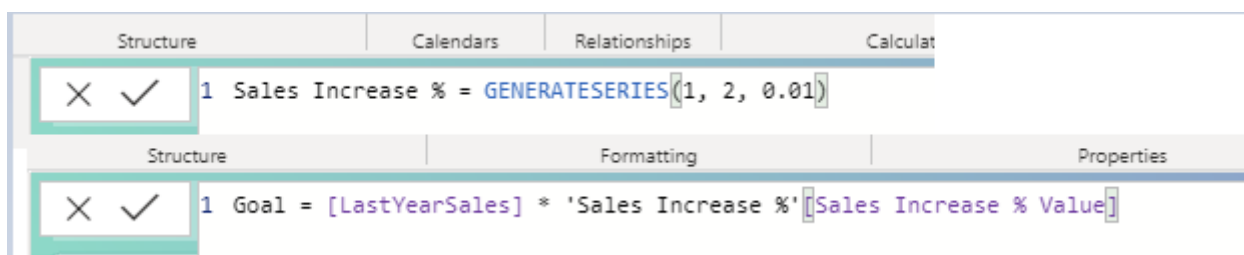


Εικόνα 4.61: Forecasting Upper-Lower

Όπου βλέπουμε γκρίζο χρώμα είναι η πρόβλεψη που έχει γίνει. Στην εικόνα 4-61 μπορούμε να δούμε την απόκλιση που μπορεί να έχει η πρόβλεψη μας γι' αυτό έχουμε upper και lower bounds.

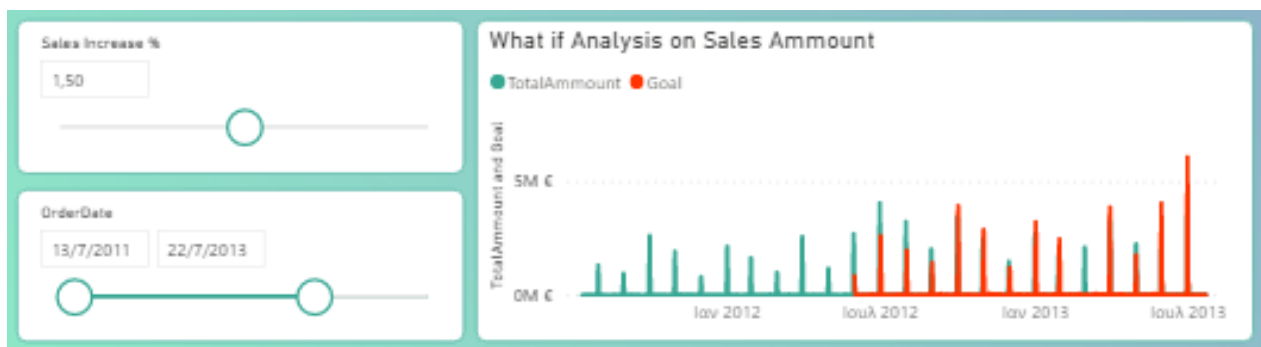
Κατασκευή What If Σεναρίου

Αυτά τα σενάρια βοηθάνε την επιχείρηση άμεσα στην λήψη αποφάσεων, καθώς της δείχνουν πως ένας δείκτης μπορεί να επηρεάσει την πορεία της. Στο συγκεκριμένο σενάριο που θα αναπτύξουμε ένα μοντέλο το οποίο θα προβλέπει πόσο απέχουμε από τον στόχο που έχουμε σαν επιχείρηση όσον αφορά τις πωλήσεις. Για την κατασκευή του θα χρησιμοποιήσουμε τις πωλήσεις του 2013 όπου κατασκευάσαμε προηγουμένως ενώ θα φτιάξουμε και δύο νέα measurement. Το ένα θα αναπαριστά έναν δεκαδικό αριθμό από το 1 έως το 2 ας τον ονομάσουμε ρυθμό ανάπτυξης. Το 1 αναπαριστά το 0% και το 2 το 100% και θα χρησιμοποιηθεί για να αναπαραστήσουμε το νέο measurement που θα υπολογίζει την επί της εκατό αύξηση του κέρδους που θέλουμε να επιτύχουμε. Άρα θα ορίζεται ως Κέρδος (επιθυμητού έτους) * ρυθμό αύξησης.

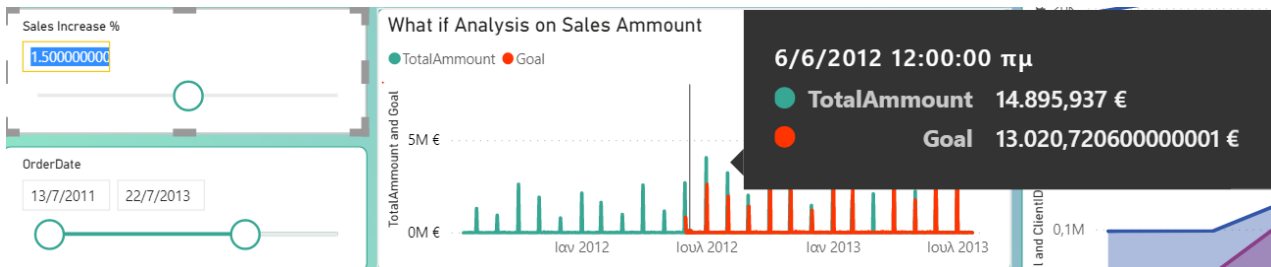


Εικόνα 4.62: Measurements για what if ανάλυση

Για τον έλεγχο της τιμής του ρυθμού αύξησης θα φτιάξουμε ένα slicer ώστε να βάζει ο χρήστης την επιθυμητή τιμή. Θα προσθέσουμε ένα ακόμα slicer ημερομηνίας με χρήση του οποίου ο χρήστης θα βλέπει το σενάριο για την επιθυμητή περίοδο. Για παράδειγμα στην εικόνα 3-61 έχουμε επιλέξει τις ημερομηνίες από 13/7/2011 έως 22/7/2013. Επομένως, θα χρησιμοποιήσει τα δεδομένα του 2012-2013 * 1.5 που έχει επιλεγθεί σαν ρυθμός αύξησης. Αυτό σημαίνει ότι ο στόχος για τις πωλήσεις μας είναι μια αύξηση 50 τα εκατό επί των κερδών του 2013. Στο διάγραμμα βλέπουμε με κόκκινο χρώμα την περίοδο υλοποιείται το σενάριο και με πράσινο τα κανονικά κέρδη της εταιρείας για το αντίστοιχο διάστημα.



Εικόνα 4.63: What if Analysis



Εικόνα 4.64: What if Analysis II

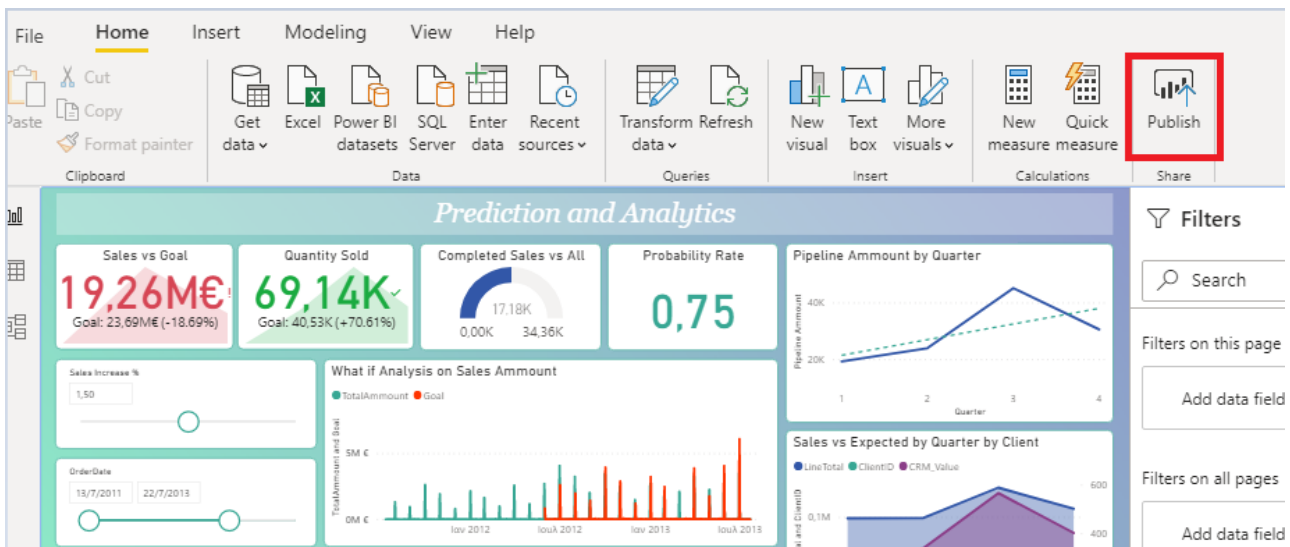
Σε αυτό το διάγραμμα μπορούμε να εφαρμόσουμε τα δύο φίλτρα που βλέπουμε στην εικόνα 4.63, το ένα είναι ένα φίλτρο χρόνου και το άλλο όπως αναφέραμε δείχνει τον ρυθμό ανάπτυξης που θέλουμε να επιτύχουμε. Συνεπώς μπορεί ο χρήστης να δει για το διάστημα και το αντίστοιχο ρυθμό ανάπτυξης που θα επιλέξει το πόσο απέχει ή όχι από το επιθυμητό αποτέλεσμα διαμορφώνοντας αντίστοιχα και την στρατηγική του.

5

Αποτελέσματα Ανάλυσης

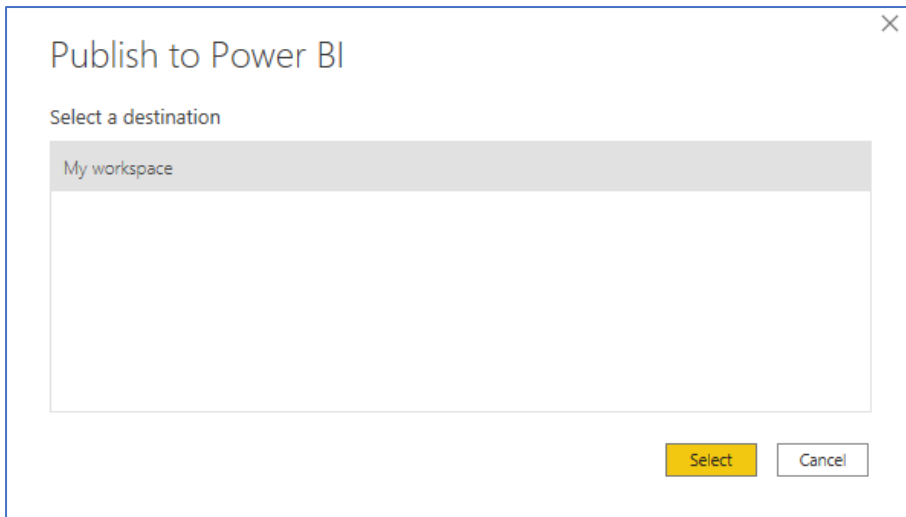
5.1 Deployment στο σύννεφο

Όταν έχουμε τελειώσει την ανάλυσή μας σειρά έχει να ανεβάσουμε τα reports μας στην cloud Υπηρεσία του Power BI (Power BI as a service). Για να γίνει αυτό πατάμε το κουμπί Publish όπως βλέπουμε στην εικόνα 5-1.

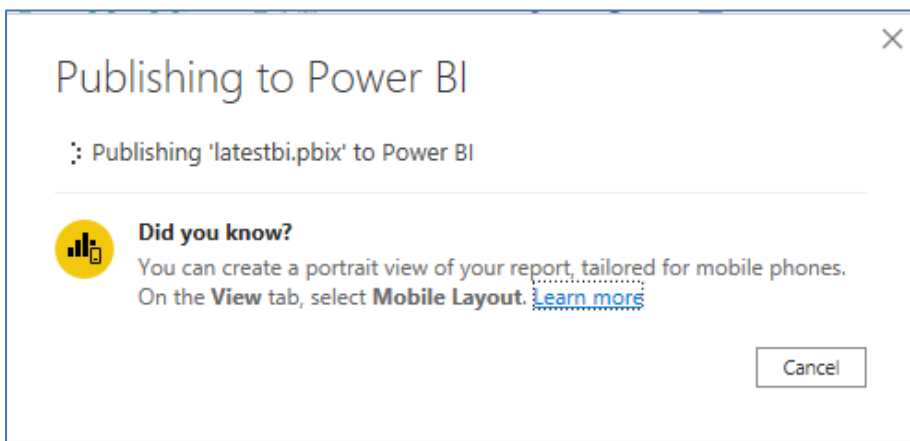


Εικόνα 5.1: Publish Report

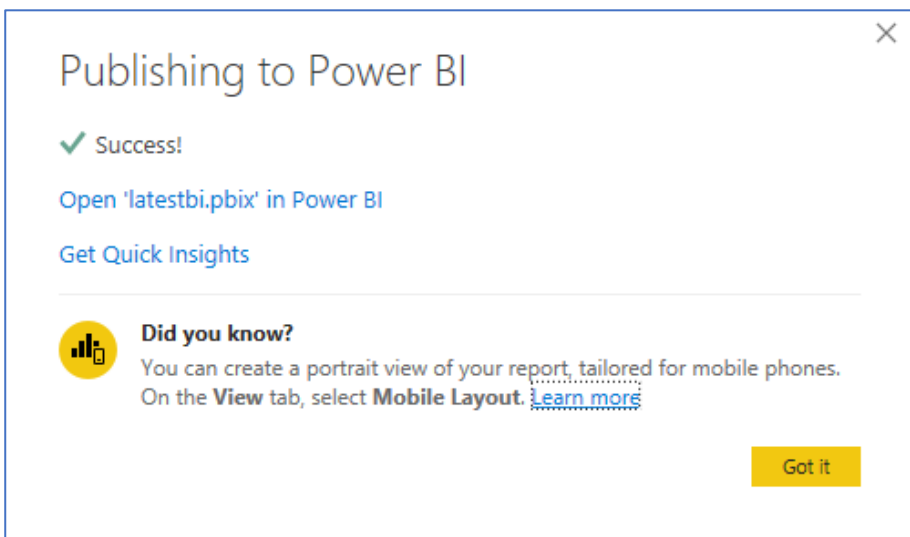
Σε αυτό το σημείο αν δεν έχουμε συνδεθεί με τον λογαριασμό μας θα μας ζητήσει τα στοιχεία μας για να συνδεθούμε. Επειδή αυτό το βήμα έχει προηγηθεί, το πρόγραμμα μας ρωτάει απευθείας σε πιο χώρο να ανεβάσει τα reports μας. Στην συγκεκριμένη περίπτωση έχει φτιαχτεί ο χώρος workspace οπότε επιλέγουμε αυτόν και μετά κάνουμε κλικ στο κουμπί **Select** (εικόνα 5-2).



Εικόνα 5.2: Επιλογή destination folder

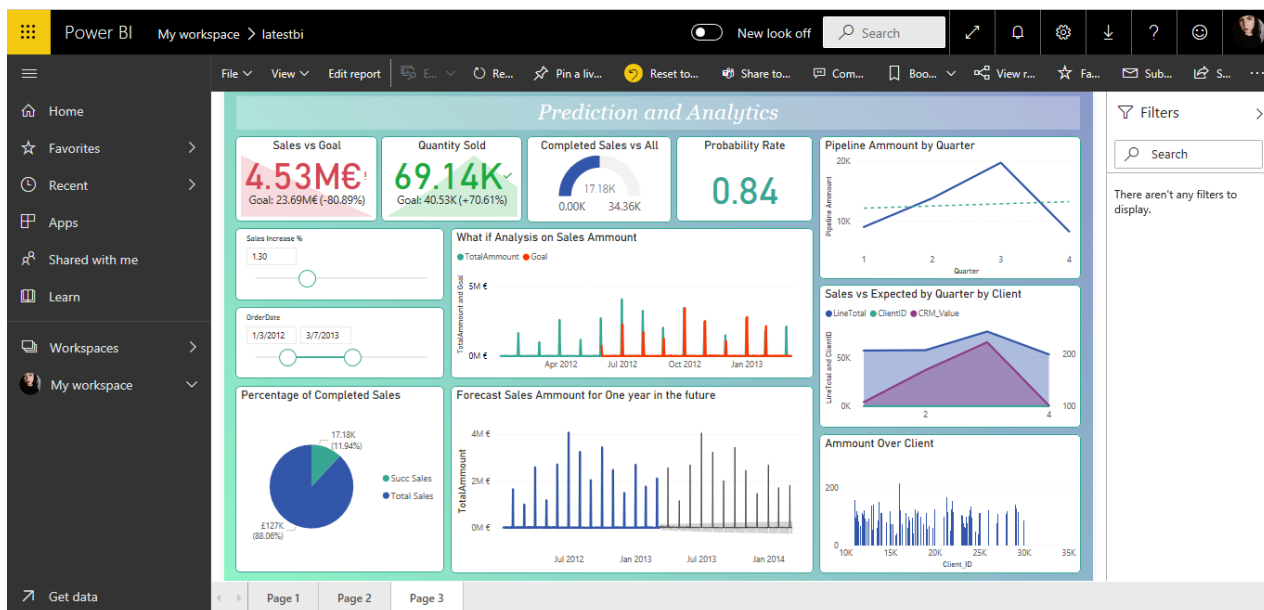


Εικόνα 5.3: Publishing to Power BI



Εικόνα 5.4: Publishing to Power BI Success

Πλέον μπορούμε να δούμε τα report μας online από οπουδήποτε απλά μπαίνοντας στον Power BI as a Service λογαριασμό μας.

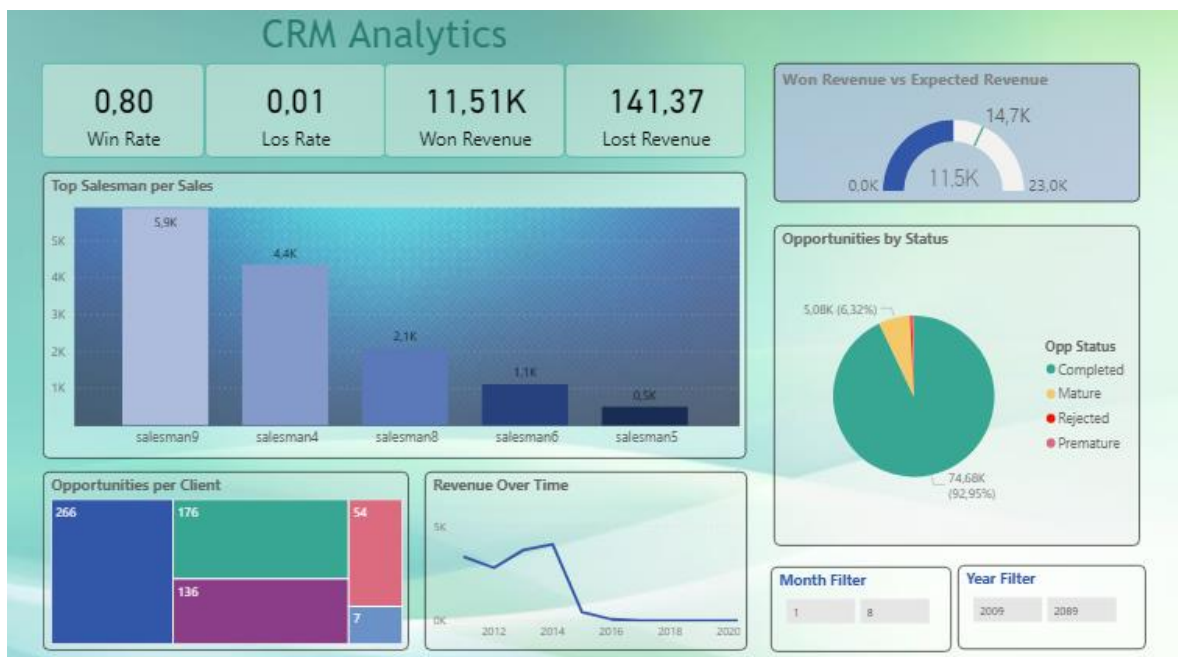


Εικόνα 5.5: Power BI Dashboard

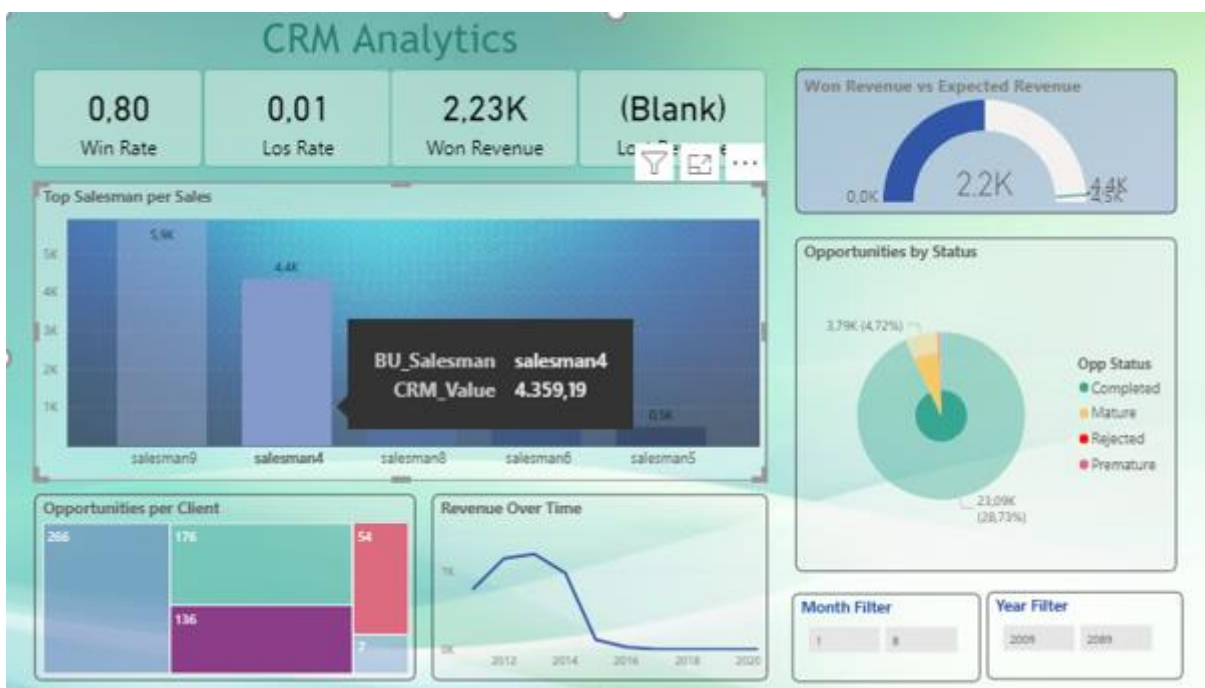
Για να μπορέσουμε πλέον να διαμοιράσουμε το report μας σε άτομα της επιχείρησης που έχουν πρόσβαση πατάμε το κουμπί share και διαμοιράζουμε το link που μας δίνει η υπηρεσία.

5.2 Διαδραστικότητα Αναφορών

Όπως έχουμε αναφέρει ένα από τα μεγαλύτερα πλεονεκτήματα που μας δίνει ένα BI εργαλείο είναι το διαδραστικό γραφικό περιβάλλον. Το γραφικό περιβάλλον είναι by default διαδραστικό χωρίς να απαιτείται κάποια ρύθμιση. Έτσι υπάρχει η δυνατότητα ο χρήστης να πατήσει πάνω σε κάποιο γράφημα αλλάζοντας το περιεχόμενο που δείχνουν τα υπόλοιπα συσχετιζόμενα γραφήματα.



Εικόνα 5.6: Αρχικό Report CRM

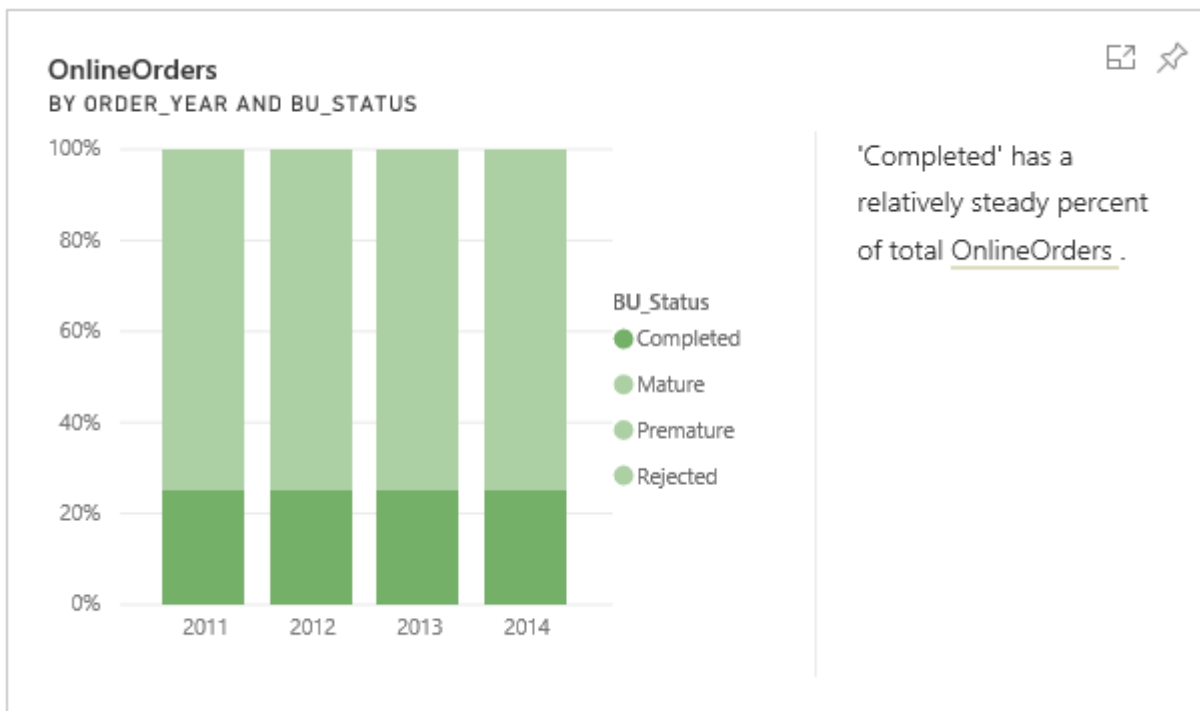


Εικόνα 5.7: Επιλογή Salesman στο Crm Report

Στην εικόνα 5-7, βλέπουμε το πως αλλάζει το παραπάνω report αν πατήσουμε πάνω στον **Salesman4** του διαγράμματος **Top Salesman per Sales**. Βλέπουμε ότι κάποια γραφήματα έχουν αλλάξει ενώ κάποια άλλα έχουν μείνει αναλλοίωτα. Για παράδειγμα ο δείκτης loss rate δεν σχετίζεται κάπως με τους πωλητές οπότε όταν έχουμε επιλέξει το διάγραμμα να δείξει πληροφορίες ως προς τον πωλητή 4 η τιμή του δείκτη αυτού θα παραμείνει ως έχει. Έτσι η επιχείρηση μπορεί να δει για παράδειγμα τι κέρδη αποφέρει ο συγκεκριμένος πωλητής και να βγάλει συμπέρασμα για το αν η πορεία του μέσα στον οργανισμό είναι καλή ή κακή. Ενώ ακόμα μπορεί να συγκρίνει με άλλες χρονικές περιόδους απλά αλλάζοντας το φίλτρο χρόνου. Στο άλλο γράφημα που βλέπουμε την τιμή Blank δεν υπήρχαν δεδομένα για το συγκεκριμένο διάστημα για τον πωλητή 4 για αυτό και είναι μηδενικό.

5.3 Insights

Το Power BI δίνει την δυνατότητα αυτόματης ανάλυσης των δεδομένων μας, ούτως ώστε να τα εξετάσει και να βρει ενδιαφέρουσες τάσεις και μοτίβα. Αυτές οι τάσεις και τα μοτίβα παρουσιάζονται με τη μορφή γραφικών που ονομάζονται Insights [19] Κάποια insights που αξίζουν να αναφερθούν είναι τα εξής:



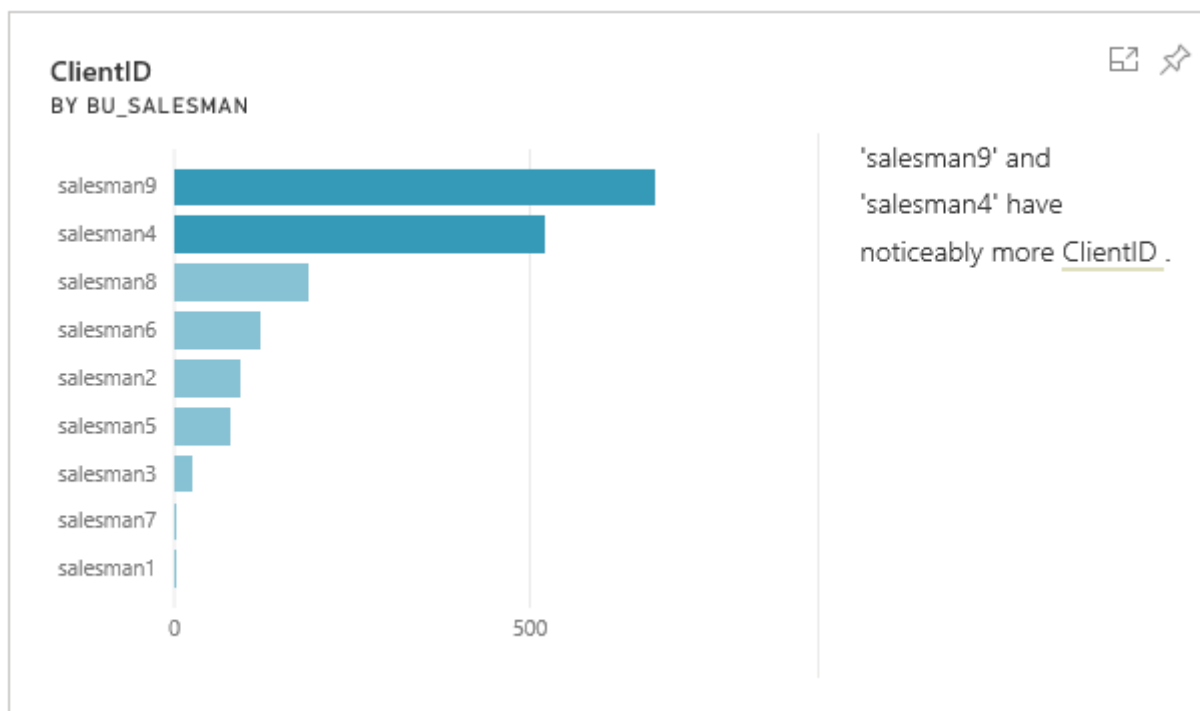
Εικόνα 5.8: Insight of online orders

Στην εικόνα 5-8 βλέπουμε ότι ένα μεγάλο ποσοστό των ευκαιριών που έγιναν πωλήσεις ήταν διαδικτυακές αγορές, επομένως διαπιστώνουμε μια προτίμηση του αγοραστικού κοινού σε αυτές. Άρα η επιχείρηση μπορεί είτε να στρέψει την στρατηγική της στις διαδικτυακές αγορές για να αυξήσει ακόμα περισσότερο το κέρδος της, είτε να προσπαθήσει να αυξήσει την διαφήμιση στις on premise αγορές στις οποίες όπως φαίνεται μειονεκτεί.



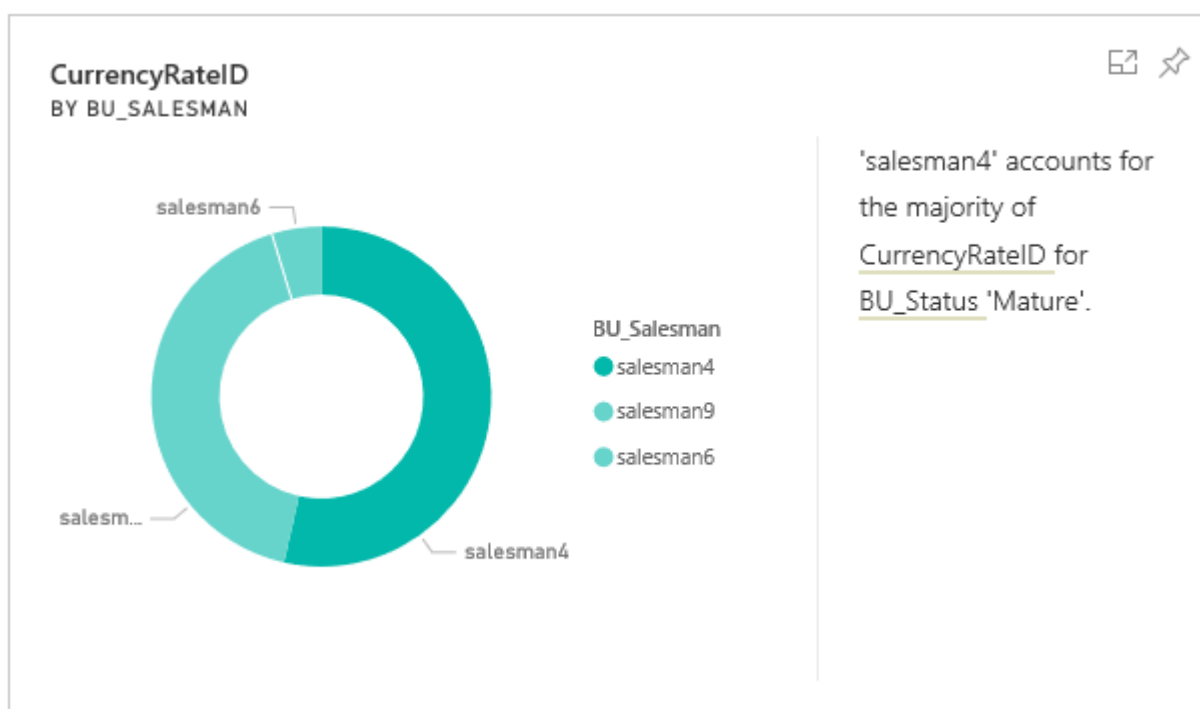
Εικόνα 5.9: Correlation of online orders

Στην εικόνα 5-9 βλέπουμε την συσχέτιση (correlation) μεταξύ των διαδικτυακών πωλήσεων και των κερδών των πωλήσεων με βάση το probability και μάλιστα παρατηρούμε ότι έχουμε μια θετική συσχέτιση (positive correlation) μεταξύ των δυο μεταβλητών, δηλαδή βλέπουμε ότι και οι δύο μεταβλητές κινούνται προς την ίδια κατεύθυνση.



Εικόνα 5.10: ClientId and Salesman

Στην εικόνα 5.10 βλέπουμε ότι οι Salesman 9 και 4 έχουν τους περισσότερους πελάτες κάτι που, αποδεικνύεται και από το αντίστοιχο διάγραμμα με τους καλύτερους πωλητές με βάση τις πωλήσεις στο οποίο και πάλι ο 9 και ο 4 ήταν οι καλύτεροι πωλητές. Άρα συμπεραίνουμε ότι οι καλύτεροι πωλητές είναι αυτοί που έχουν τους περισσότερους πελάτες.



Εικόνα 5.11: Salesman and Status

Τέλος, βλέπουμε ότι μεγάλο μέρος των ευκαιριών που δεν έχουν εξελιχθεί ούτε σε πωλήσεις αλλά ούτε έχουν απορριφθεί ανήκαν στον Salesman⁴. Άρα προκύπτει το συμπέρασμα ότι ενώ ο συγκεκριμένος πωλητής είναι από τους καλύτερους που έχει η εταιρεία έχει και πολλές ευκαιρίες τις οποίες δεν καταφέρνει να γυρίσει σε πωλήσεις.

6

Συμπεράσματα

Η συγκεκριμένη διπλωματική εργασία είχε ως στόχο να αναδείξει τα οφέλη τα οποία μπορεί να έχει κάποιος οργανισμός, εάν υιοθετήσει τα νέα και καινοτόμα επιχειρηματικά συστήματα Business Intelligence και Analytics χρησιμοποιώντας δεδομένα ERP και CRM της επιχείρησης. Συμπεραίνουμε ότι είναι σημαντικό κάθε επιχείρηση να αντιμετωπίζει τα δεδομένα της ως ένα πολύτιμο αγαθό διότι μέσα από αυτά μπορεί να εξάγει χρήσιμες πληροφορίες οι οποίες μετά την εφαρμογή διαδικασιών αναλυτικής οδηγούν σε χρήσιμα συμπεράσματα για τον τρόπο λειτουργίας της επιχείρησης. Εξίσου σημαντική είναι και η διαχείριση αυτών διότι οποιαδήποτε αλλοίωση στα δεδομένα μπορεί να οδηγήσει σε εξαγωγή λάθος συμπερασμάτων. Τα συστήματα επιχειρησιακής ευφυΐας αυτά βοηθούν ιδιαίτερα στις στρατηγικές σχετικά με τη λήψη αποφάσεων και εξοβελίζουν τον παράγοντα της τύχης. Διαδέχονται τα DSS διότι προσφέρουν πολλά περισσότερα από απλή υποστήριξη αποφάσεων. Δίνουν δυνατότητες όπως είδαμε εφαρμογής πιθανών σεναρίων μέσω των οποίων μπορούμε να διακρίνουμε στρατηγικές που πρέπει να ακολουθηθούν στο άμεσο μέλλον. Επιπλέον, προτείνεται μία ολοκληρωμένη μεθοδολογία υλοποίησης έργων Επιχειρησιακής Ευφυΐας, η οποία δύναται να έχει εφαρμογή μέσα σε κάποιον οργανισμό ή κάποια επιχείρηση. Έπειτα από μελέτη του θεωρητικού πλαισίου στο χώρο του BI σχεδιάστηκαν τα βήματα αυτής της μεθοδολογίας. Δόθηκε ιδιαίτερη βάση στον καθορισμό ξεκάθαρων σταδίων για την υλοποίηση του έργου. Στην πορεία της εργασίας χρησιμοποιήθηκε ένα υποθετικό σενάριο, ώστε να εφαρμοστούν όλα τα βήματα σε πρακτικό επίπεδο. Σε πρώτο επίπεδο αναλύθηκαν οι απαιτήσεις του προαναφερθέντος σεναρίου, έτσι ώστε να τεθεί ορθά η στοχοθεσία της υλοποίησης. Σε επόμενο βήμα έγινε η απαραίτητη διαδικασία ETL, μέσω της οποίας αντλούμε δεδομένα από συστήματα των πελατών (ERP, CRM) και κάνουμε μεταφορά στην βάση δεδομένων που δημιουργήσαμε. Με τον τρόπο αυτό, δομούμε την πληροφορία όπως θεωρούμε

σκοπίμο. Στη συνέχεια έγινε αναλυτική επεξεργασία στα δεδομένα, και δημιουργήθηκαν τα τελικά reports. Γενικότερα δεν κρίνεται ιδιαίτερα απαιτητική διαδικασία η υλοποίηση της μεθοδολογίας BI, για τη δόμηση ενός ολοκληρωμένου συστήματος, το οποίο εκπροσωπεί τις ανάγκες ενός οργανισμού. Ωστόσο, ο παράγοντας της δυσκολίας ή μη, έχει άμεση εξάρτηση από τη διαθεσιμότητα της πληροφορίας για την εισαγωγή της στο σύστημα. Η πληροφορία αυτή προέρχεται ιδίως από ERP και CRM συστήματα. Όσο ορθότερη και πιο δομημένη είναι η εισαχθείσα πληροφορία, τόσο μεγαλύτερη είναι και η ανάλυση η οποία προσφέρουμε στον τελικό χρήστη, χωρίς να χρειαστεί επιπλέον μόχθος για εκκαθάριση δεδομένων. Στον αντίποδα, εφόσον η πληροφορία είναι διάσπαρτη, ανομοιογενής ή έχει ελλείμματα, η δυσκολία ακολουθεί αύξουσα πορεία. Η υιοθέτηση της λύσης αυτής προσφέρει θετικά αποτελέσματα και οφέλη όπως είναι η βελτίωση της λήψης αποφάσεων και η βελτίωση σε εσωτερικές επιχειρηματικές διαδικασίες. Στόχο αποτελεί η εύρεση νέων πηγών εσόδων και ως εκ τούτου το κέρδος του ανταγωνιστικού πλεονεκτήματος σε σχέση με τις υπόλοιπες επιχειρήσεις του κλάδου. Τα συστήματα BI είναι ωφέλιμα στις εταιρείες, αφού είναι βοηθητικά στον εντοπισμό των τάσεων της αγοράς και του ρίσκου που είναι πιθανό να έρθουν αντιμέτωπες ανάλογα με τα βήματα και τις κινήσεις που ακολουθούν. Φυσικά όπως συμβαίνει με όλους τους τομείς, οι οποίοι παρουσιάζουν μειονεκτήματα, έτσι και τα συστήματα BI εμφανίζουν μερικά τα οποία εστιάζονται πρωτίστως στο πολύ υψηλό κόστος που απαιτείται για την εγκατάστασή τους. Εκτός αυτού, είναι εξίσου υψηλό το κόστος αποθήκευσης δεδομένων προηγούμενων ετών, όπως και αυτό της εκπαίδευσης και εξειδίκευσης του προσωπικού. Επίσης, ενδέχεται ο τελικός χρήστης να λάβει εσφαλμένες πληροφορίες, αν τα δεδομένα τα οποία δέχεται το σύστημα έχουν ελλείψεις και δεν είναι δομημένα ορθά. Όπως προκύπτει από αυτά, είναι πάντα απαραίτητος ο έλεγχος, όταν τα δεδομένα υφίστανται ανανέωση στο σύστημα, ώστε να κατοχυρώνεται η βεβαιότητα του τελικού χρήστη ως προς την αξιοπιστία των αναφορών που βλέπει. Τέλος, αφού τα συστήματα Business Intelligence και Analytics είναι ένας νέος κλάδος στην αγορά εργασίας, σίγουρα προβλέπονται βελτιώσεις εντός των επόμενων ετών, οι οποίες θα βοηθήσουν στην αύξηση της αξιοπιστίας τους, αλλά και στο κομμάτι της λειτουργικότητάς τους.

Βιβλιογραφία

- [1] Keith D. Foote. (2017) A Brief History of Business Intelligence <https://www.dataversity.net/brief-history-business-intelligence/#>
- [2] Oracle (2020) <https://www.oracle.com/database/what-is-data-management/>
- [3] Deborah Henderson, Susan Earley. (2017) DAMA-DMBOK Data Management Body of Knowledge 2nd Edition
- [4] Wikipedia (2020) Star Schema https://en.wikipedia.org/wiki/Star_schema
- [5] Σημειώσεις του μαθήματος Συστήματα Υποστήριξης Αποφάσεων
- [6] Software Testing Help (2020) Top Data Warehouse Tools <https://www.softwaretestinghelp.com/data-warehouse-tools/>
- [7] PwC Academy (2019) Database concepts-Relational, NoSQL and Data Warehouse pdf
- [8] Arvanitis S., Loukis E. (2009). Information and Communication Technologies, Human Capital, Workplace Organization and Labour Productivity in Greece and Switzerland: A Comparative Study Based on Firm-level Data, Information Economics and Policy (Social Sciences Citation Index, ACM Digital Library)
- [9] PwC Academy (2019) Business Analytics pdf
- [10] https://el.wikipedia.org/wiki/%CE%A5%CF%80%CE%BF%CE%BB%CE%BF%CE%B3%CE%B9%CF%83%CF%84%CE%B9%CE%BA%CF%8C_%CE%BD%CE%AD%CF%86%CE%BF%CF%82
- [11] (2017) <https://www.business.qld.gov.au/running-business/it/cloud-computing/benefits>
- [12] (2020) <https://www.atscale.com/blog/what-is-a-cloud-data-warehouse/>
- [13] Hussain Al-Aqrabi Lu Liu Richard Hill Nick Antonopoulos (2015) Cloud BI: Future of business intelligence in the Cloud
- [14] <https://docs.microsoft.com/en-us/azure/data-factory/#:~:text=Azure%20Data%20Factory%20is%20Azure's,with%20full%20compatibility%20in%20ADF.>
- [15] <https://azure.microsoft.com/en-us/solutions/data-lake/>
- [16] <https://www.saviantconsulting.com/blog/10-reasons-why-choose-microsoft-azure.aspx>
- [17] <https://fileinfo.com/extension/dacpac#:~:text=What%20is%20a%20DACPAC%20file,Server%20databases%20and%20vendor%20applications.>
- [18] <https://docs.microsoft.com/en-us/azure/data-factory/quickstart-create-data-factory-portal#create-a-linked-service>

[19]<https://docs.microsoft.com/en-us/power-bi/consumer/end-user-insight-types#:~:text=You%20can%20ask%20Power%20BI,visuals%20that%20are%20called%20Insights.>

Παράρτημα I Κώδικας σε sql

1) Κώδικας για την κατασκευή της όψης πωλήσεων

```

1. SET QUOTED_IDENTIFIER ON
2. GO
3.
4.
5.
6. CREATE View[dbo].[Sales_View]
7.
8. as
9.
10.
11.
12. SELECT Hd.[SalesOrderID]
13. , cast (LEFT (Hd. [SalesOrderID],4) asint) asGeoID
14. ,Hd.[OrderDate]
15. ,Hd.[DueDate]
16. ,Hd.[ShipDate]
17. ,Hd.[Status]
18. ,Hd.[OnlineOrderFlag]
19. ,Hd.[SalesOrderNumber]
20. ,Hd.[PurchaseOrderNumber]
21. ,Hd.[AccountNumber]
22. ,Hd.[CustomerID]
23. ,Hd.[SalesPersonID]
24. ,Hd.[TerritoryID]
25. ,Hd.[BillToAddressID]
26. ,Hd.[ShipToAddressID]
27. ,Hd.[ShipMethodID]
28. ,Hd.[CreditCardID]
29. ,Hd.[CreditCardApprovalCode]
30. ,Hd.[CurrencyRateID]
31. --,Hd.[SubTotal]
32. --,Hd.[TaxAmt]
33. --,Hd.[Freight]
34. --,Hd.[TotalDue]
35. ,Hd.[Comment]
36. --,Hd.[rowguid]
37. --,Hd.[ModifiedDate]
38. ,Det.[SalesOrderDetailID]
39. ,Det.[CarrierTrackingNumber]
40. ,Det.[OrderQty]
41. ,Det.[ProductID]
42. ,Det.[SpecialOfferID]
43. ,Det.[UnitPrice]

```

```

44. ,Det.[UnitPriceDiscount]
45. ,Det.[LineTotal]
46. ,Det.[rowguid]
47. ,Det.[ModifiedDate]
48. FROM [S1_SalesOrderHeader]Hd
49. INNER JOIN [dbo].[S1_SalesOrderDetail]Det
50. on Det.SalesOrderID=Hd.SalesOrderID
51. GO

```

2) Κώδικας για τον πίνακα CRMBreakdown

```

1. SET ANSI_NULLS ON
2. GO
3.
4. SET QUOTED_IDENTIFIER ON
5. GO
6.
7. CREATE TABLE [dbo].[CRMDData_Breakdown] (
8.     [OppID] [smallint] NULL,
9.     [StartDate] [date] NULL,
10.     [Enddate] [date] NULL,
11.     [Qty] [float] NULL,
12.     [Daily_Qty] [float] NULL,
13.     [ref_date] [date] NULL
14. ) ON [PRIMARY]
15. GO

```

3) Κώδικας για την όψη CRMFinal

```

1. SET ANSI_NULLS ON
2. GO
3.
4. SET QUOTED_IDENTIFIER ON
5. GO
6.
7. CREATE View[dbo].[CRM_Final]as
8.
9. SELECT     Data.[OppID]
10.    ,Data.[StartDate]
11.    ,Data.[Enddate]
12.    ,Data.[Qty]
13.    ,Data.[Daily_Qty]asCRM_Value
14.    ,Data.[ref_date],
15.        (select top 1[Opp BU]from crm wherecrm.OppID=Data.OppID)as BU,
16.        (select top 1[OppStatus]from crm
17.        wherecrm.OppID=Data.OppID)asBU_Status,
18.        (select top 1[OppName]from crm wherecrm.OppID=Data.OppID)as
19.        BU_Name,
20.        (select top 1[OppSalesman]from crm wherecrm.OppID=Data.OppID)as
21.        BU_Salesman,
22.        (select top 1ClientIDfrom crm
23.        wherecrm.OppID=Data.OppID)asClientID
24. FROM [dbo].[CRMDData_Breakdown]Data
25. GO

```

4) Κώδικας για το SalesPipeline

```

1. SET ANSI_NULLS ON
2. GO
3.
4. SET QUOTED_IDENTIFIER ON
5. GO
6.
7. CREATE TABLE [dbo].[QuartOppSales] (
8.     [OppID] [smallint] NULL,
9.     [StartDate] [date] NULL,
10.    [Enddate] [date] NULL,
11.    [Qty] [float] NULL,
12.    [CRM_Value] [float] NULL,
13.    [BU] [nvarchar] (255) NULL,
14.    [BU_Status] [nvarchar] (255) NULL,
15.    [BU_Name] [nvarchar] (255) NULL,
16.    [BU_Salesman] [nvarchar] (255) NULL,
17.    [ClientID] [float] NULL,
18.    [Year] [int] NULL,
19.    [Quarter] [int] NULL,
20.    [QStartDate] [date] NULL,
21.    [QEndDate] [date] NULL,
22.    [DaysInQuarter] [int] NULL,
23.    [SalesOrderID] [int] NOT NULL,
24.    [GeoID] [int] NULL,
25.    [OrderDate] [datetime] NOT NULL,
26.    [DueDate] [datetime] NOT NULL,
27.    [ShipDate] [datetime] NULL,
28.    [Status] [tinyint] NOT NULL,
29.    [OnlineOrderFlag] [bit] NOT NULL,
30.    [SalesOrderNumber] [nvarchar] (25) NOT NULL,
31.    [PurchaseOrderNumber] [nvarchar] (25) NULL,
32.    [AccountNumber] [nvarchar] (15) NULL,
33.    [CustomerID] [int] NOT NULL,
34.    [SalesPersonID] [int] NULL,
35.    [TerritoryID] [int] NULL,
36.    [BillToAddressID] [int] NOT NULL,
37.    [ShipToAddressID] [int] NOT NULL,
38.    [ShipMethodID] [int] NOT NULL,
39.    [CreditCardID] [int] NULL,
40.    [CreditCardApprovalCode] [varchar] (15) NULL,
41.    [CurrencyRateID] [int] NULL,
42.    [Comment] [nvarchar] (128) NULL,
43.    [SalesOrderDetailID] [int] NOT NULL,
44.    [CarrierTrackingNumber] [nvarchar] (25) NULL,
45.    [OrderQty] [smallint] NOT NULL,
46.    [ProductID] [int] NOT NULL,
47.    [SpecialOfferID] [int] NOT NULL,
48.    [UnitPrice] [money] NOT NULL,

```

```
49.         [UnitPriceDiscount] [money] NOT NULL,  
50.         [LineTotal] [numeric] (38,6) NOT NULL,  
51.         [rowguid] [uniqueidentifier] NOT NULL,  
52.         [ModifiedDate] [datetime] NOT NULL,  
53.         [DateDistance] [int] NULL  
54.     ) ON [PRIMARY]  
55. GO
```