

# מסמך אפיון BI SOLUTION

**Date: 21/11/2022**

**Version: 1.00**

**Written by: Arbel Ziskind**

# תוכן עניינים

1.	כללי.....	3
1.1	מטרת הפרויקט.....	3
2.	אפיון טכני.....	4
2.1	ארכיטקטורת הפתרון.....	4
2.2	דרישות קדם.....	5
3.	אפיון פונקציונאלי.....	6
3.1	ארכיטקטורת הפתרון.....	6
3.1.1	– מסמך S2T ותרשימי ERD.....	6
3.1.2	– תיאור תהליכי ה-ETL.....	8
3.1.3	– תיאור טבלאות הנתונים ב-Datawarehouse.....	15
3.1.4	– תיאור המדדים, היררכיות, תרשימים והטבלאות.....	15
4.	תוכנית עבודה.....	23

## 1. כללי

### 1.1 מטרת הפרויקט

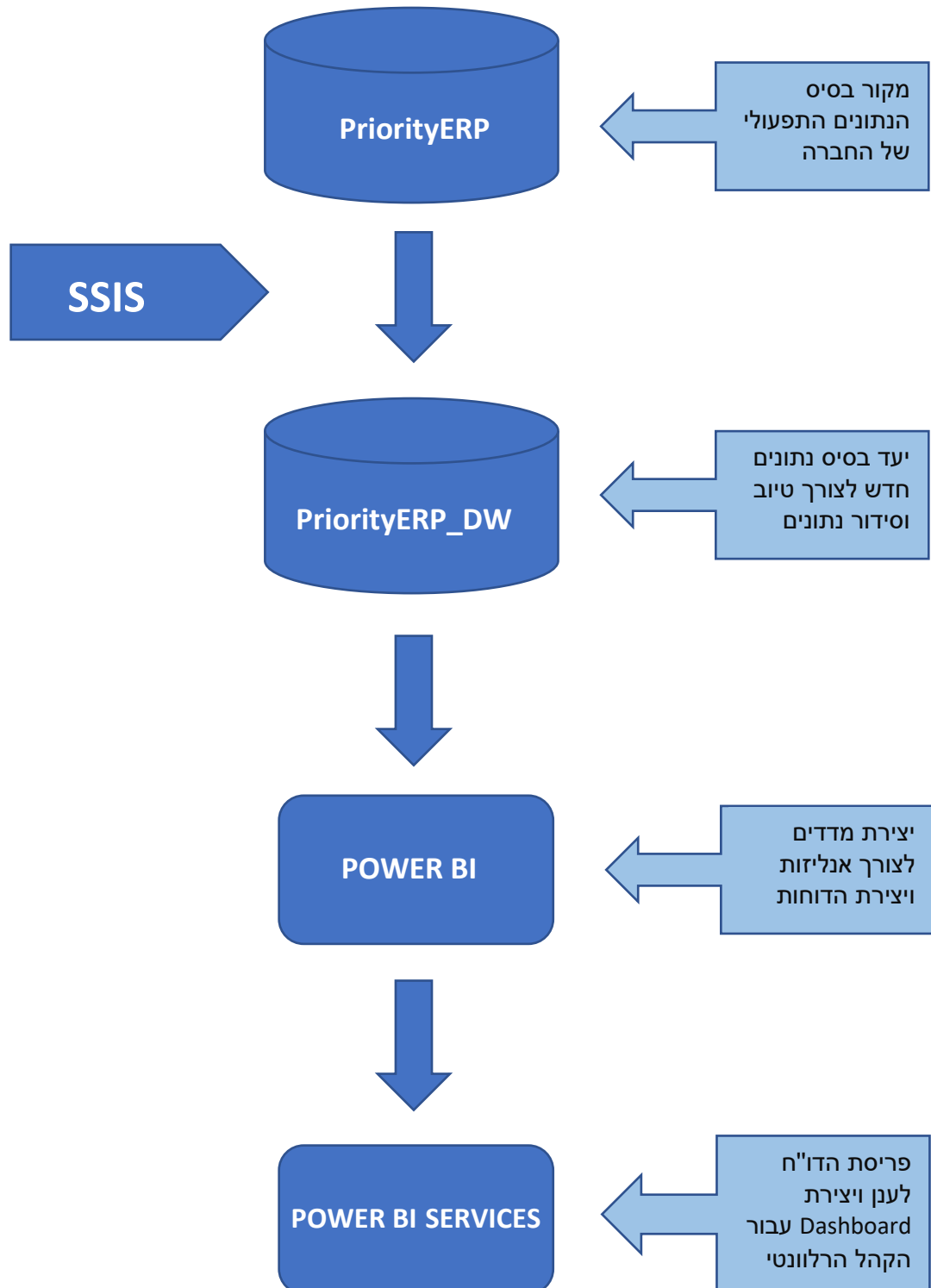
מטרת הפרויקט היא בניית פתרון BI עבור עסק המוכר כלי נגינה שמטרתו היא הצגת הנתונים לצורך מקסום הרווחים ו/או חסכון בעלויות.

יעדים :

- יצירת בסיס נתונים (Data Warehouse) המאגד את הנתונים הרלוונטיים עבור האנליזות תוך סינון וארגון מחדש של המידע.
- משיכת הנתונים מבסיס הנתונים שנוצר ויצירת ממשק גרפי המספק תובנות ומסקנות עבור הארגון על פעילותו השוטפת.

## 2. אפיון טכני

### 2.1 ארכיטקטורת הפתרון



## תיאור תמציתי של ארכיטקטורת הפתרון

- תהליך ה-ETL יבוצע באמצעות Visual Studio – SSIS ותפקידו למשוך את הנתונים מבסיס המידע התפעולי של החברה (PriorityERP) ולהעבירו לבסיס מידע Data Warehouse חדש שיצרנו (PriorityERP\_DW) תוך ארגון וסידור הנתונים הרלוונטיים לצורך ביצוע האנליזות ובניית הדוחות.
- תהליך יצירת המודל, המדדים והדוחות לצורך האנליזות ויצירת הויזואליזציה בשכבה הסמנטית יבוצע בעזרת ה-POWER BI.
- פריסת המודל והדוחות לענן (POWER BI SERVICE) ויצירת Dashboard לאנשי הארגון והגדרת עדכון המידע באמצעות תוסף POWER BI GATEWAY.

## 2.2 דרישות קדם

- חיבור פעיל לשרת SQL.
- גישה למאגר מידע תפעולי (PriorityERP)
- תוכנת VisualStudio עם תוסף SSIS.
- תוכנת SSMS.
- תוכנת Power BI Desktop.
- חשבון ב-Microsoft ומייל ארגוני עבור פריסת הדוחות לענן.

## 3. אפיון פונקציונאלי

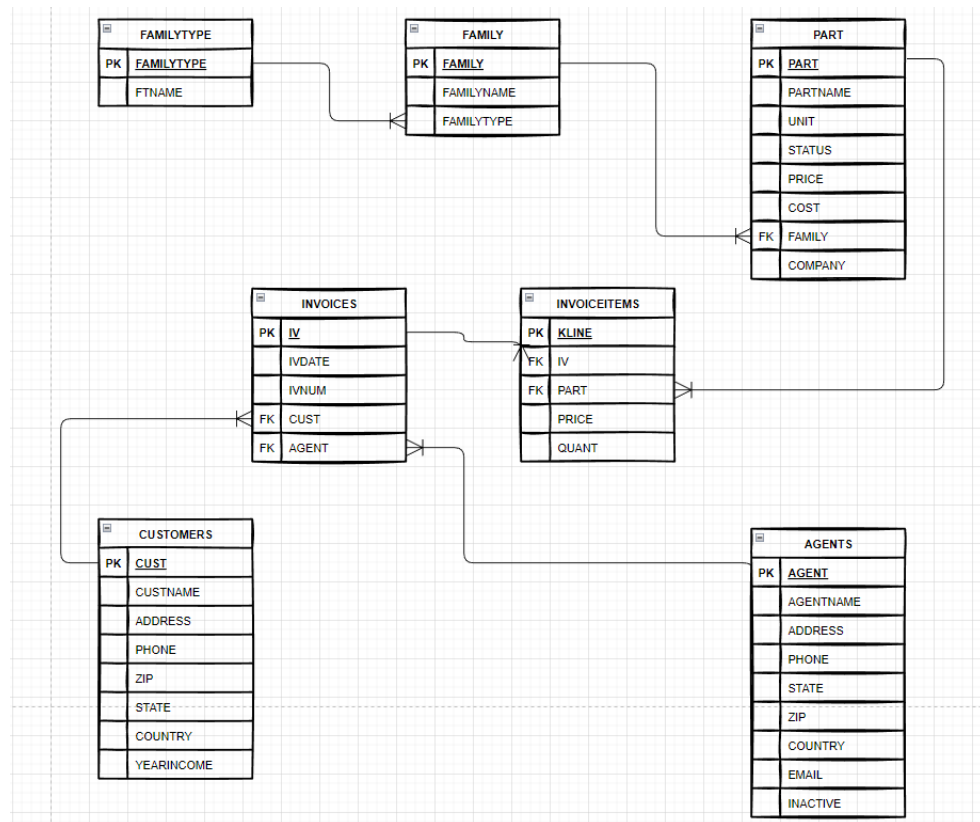
### 3.1 ארכיטקטורת הפתרון

#### 3.1.1 – מסמך S2T ותרשימי ERD.

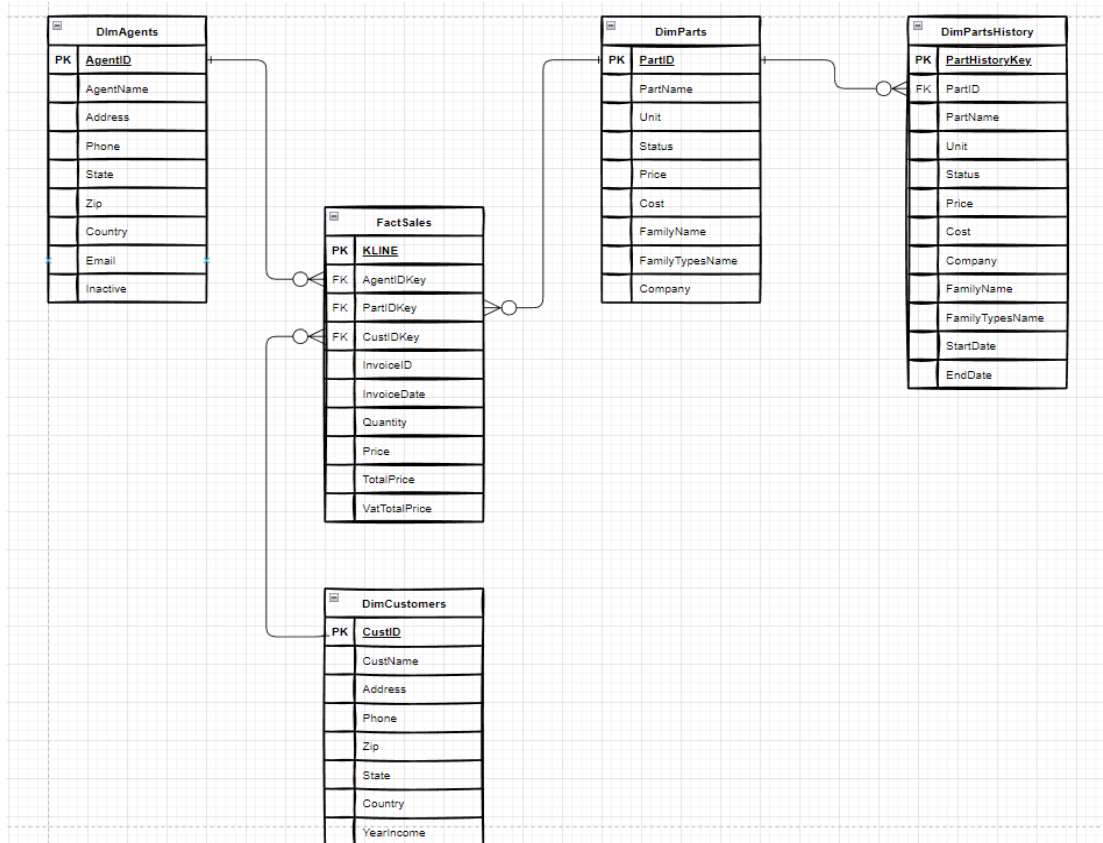
**S2T:**

Source Database	Source Table	Source Column	Source Type	Transformation	Target Database	Target Table	Target Column	Index	Target Type	Notes
PriorityERP	PART	PART	INT		PriorityERP_DW	DimParts	PartID		INT	
PriorityERP	PART	PARTNAME	NVARCHAR		PriorityERP_DW	DimParts	PartName		NVARCHAR	
PriorityERP	PART	UNIT	INT		PriorityERP_DW	DimParts	Unit		INT	
PriorityERP	PART	STATUS	NCHAR		PriorityERP_DW	DimParts	Status		NCHAR	
PriorityERP	PART	PRICE	INT		PriorityERP_DW	DimParts	Price		INT	
PriorityERP	PART	COST	INT		PriorityERP_DW	DimParts	Cost		INT	
PriorityERP	PART	FAMILY	INT		PriorityERP_DW	DimParts	Family		INT	
PriorityERP	PART	COMPANY	NVARCHAR		PriorityERP_DW	DimParts	Company		NVARCHAR	
PriorityERP	FAMILY	FAMILYNAME	NVARCHAR	SELECT PART, PARTNAME, UNIT, STATUS, PRICE, COST, FAMILYNAME, FTNAME FROM PART p JOIN FAMILY f ON p.FAMILY = f.FAMILY JOIN FAMILYTYPES ft ON f.FAMILYTYPE = ft.FAMILYTYPE	PriorityERP_DW	DimParts	FAMILYNAME		NVARCHAR	
PriorityERP	FAMILYTYPE	FTNAME	NVARCHAR		PriorityERP_DW	DimParts	FTNAME		NVARCHAR	
PriorityERP	CUSTOMERS	CUST	INT		PriorityERP_DW	DimCustomers	CustID		INT	
PriorityERP	CUSTOMERS	CUSTNAME	NVARCHAR		PriorityERP_DW	DimCustomers	CustName		NVARCHAR	
PriorityERP	CUSTOMERS	ADDRESS	NVARCHAR		PriorityERP_DW	DimCustomers	Address		NVARCHAR	
PriorityERP	CUSTOMERS	PHONE	NVARCHAR		PriorityERP_DW	DimCustomers	Phone		NVARCHAR	
PriorityERP	CUSTOMERS	ZIP	NVARCHAR		PriorityERP_DW	DimCustomers	Zip		NVARCHAR	
PriorityERP	CUSTOMERS	STATE	NVARCHAR		PriorityERP_DW	DimCustomers	State		NVARCHAR	
PriorityERP	CUSTOMERS	COUNTRY	NVARCHAR		PriorityERP_DW	DimCustomers	Country		NVARCHAR	
PriorityERP	CUSTOMERS	YEARINCOME	INT		PriorityERP_DW	DimCustomers	YearIncome		INT	
PriorityERP	AGENTS	AGENT	INT		PriorityERP_DW	DimAgents	AgentID		INT	
PriorityERP	AGENTS	AGENTNAME	NVARCHAR		PriorityERP_DW	DimAgents	AgentName		NVARCHAR	
PriorityERP	AGENTS	ADDRESS	NVARCHAR		PriorityERP_DW	DimAgents	Address		NVARCHAR	
PriorityERP	AGENTS	PHONE	NVARCHAR		PriorityERP_DW	DimAgents	Phone		NVARCHAR	
PriorityERP	AGENTS	STATE	NVARCHAR		PriorityERP_DW	DimAgents	State		NVARCHAR	
PriorityERP	AGENTS	ZIP	NVARCHAR		PriorityERP_DW	DimAgents	Zip		NVARCHAR	
PriorityERP	AGENTS	COUNTRY	NVARCHAR		PriorityERP_DW	DimAgents	Country		NVARCHAR	
PriorityERP	AGENTS	EMAIL	NVARCHAR		PriorityERP_DW	DimAgents	Email		NVARCHAR	
PriorityERP	AGENTS	INACTIVE	BIT		PriorityERP_DW	DimAgents	Inactive		BIT	
PriorityERP	INVOICES	IV	INT		PriorityERP_DW	FactSales	InvoiceID		INT	
PriorityERP	INVOICES	IVDATE	DATETIME		PriorityERP_DW	FactSales	InvoiceDate		DATETIME	
PriorityERP	INVOICES	IVNUM	NVARCHAR							
PriorityERP	INVOICES	CUST	INT		PriorityERP_DW	FactSales	CustIDKey		INT	
PriorityERP	INVOICES	AGENT	INT		PriorityERP_DW	FactSales	AgentIDKey		INT	
PriorityERP	INVOICEITEMS	KLINE	INT	SELECT IV, IVDATE, CUST, AGENT, KLINE, PART, PRICE, QUANT, PRICE * QUANT, PRICE * WUANT * 0.17 FROM INVOICES I JOIN INVOICEITEMS II ON I.IV = II.IV	PriorityERP_DW	FactSales	KLine		INT	
PriorityERP	INVOICEITEMS	IV	INT		PriorityERP_DW	FactSales	Invoice		INT	
PriorityERP	INVOICEITEMS	PART	INT		PriorityERP_DW	FactSales	PartIDKey		INT	
PriorityERP	INVOICEITEMS	PRICE	INT		PriorityERP_DW	FactSales	Price		INT	
PriorityERP	INVOICEITEMS	QUANT	INT		PriorityERP_DW	FactSales	Quantity		INT	
PriorityERP	INVOICEITEMS	PRICE, QUANT	INT, INT	PRICE * QUANT	PriorityERP_DW	FactSales	TotalPrice		INT	
PriorityERP	INVOICEITEMS	PRICE, QUANT	INT, INT	PRICE * QUANT * 0.17	PriorityERP_DW	FactSales	VatTotalPrice		FLOAT	

## ERD של בסיס המידע התפעולי:

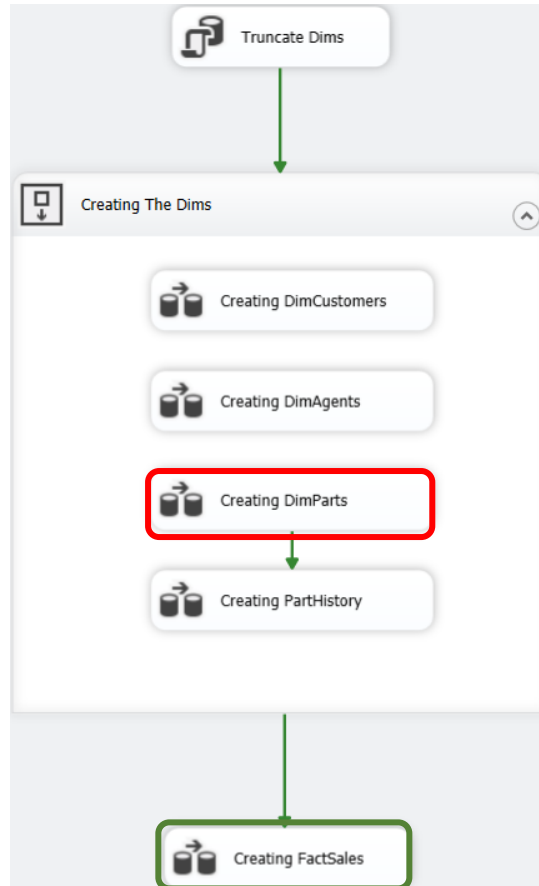


## ERD של ה-Datwarehouse:

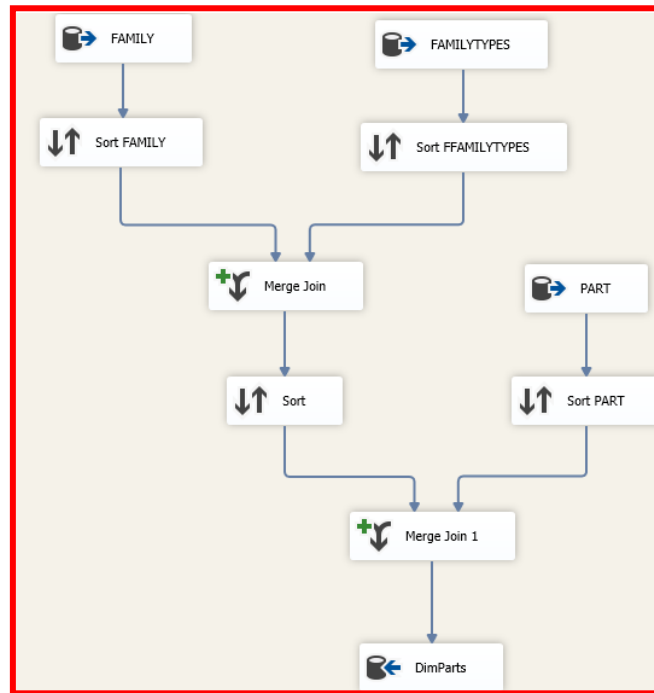


### 3.1.2 – תיאור תהליכי ה-ETL

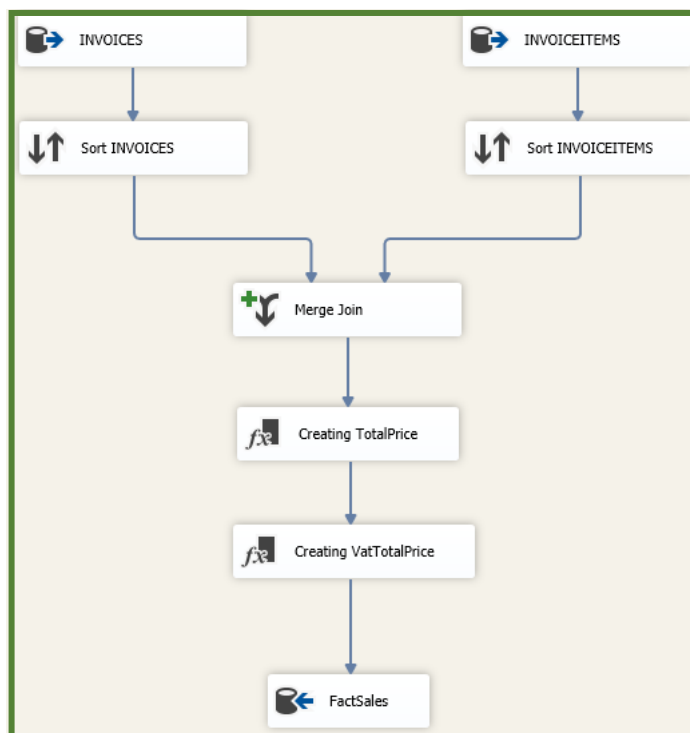
- יצירת חבילה בשם OneTime שתפקידה ליצור ולאכלס במידע את טבלאות ה-Dim וה-Fact. לצורך יצירת DimParts נבצע JOIN עם טבלאות FAMILY ו-FAMILYTYPES ולאחר מכן ניצור ממנה את טבלת ההסטוריה DimPartsHistory.





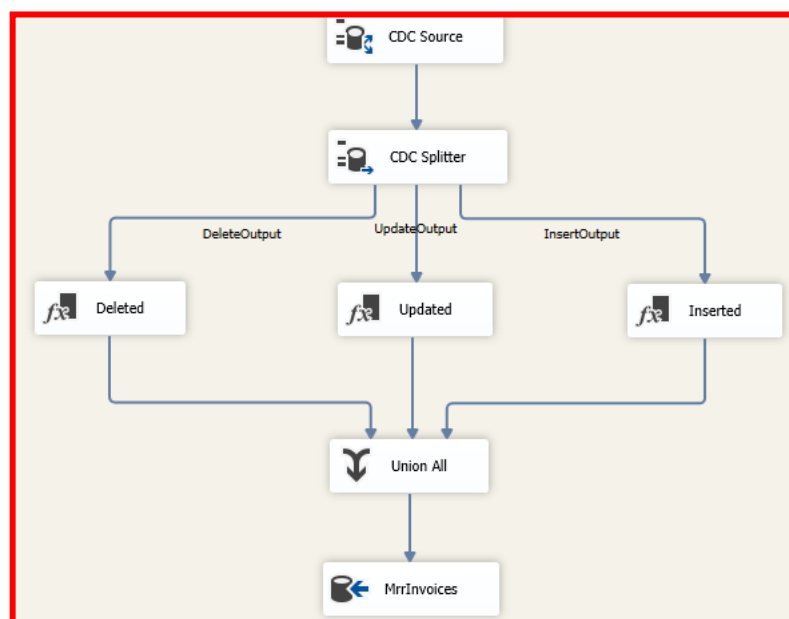


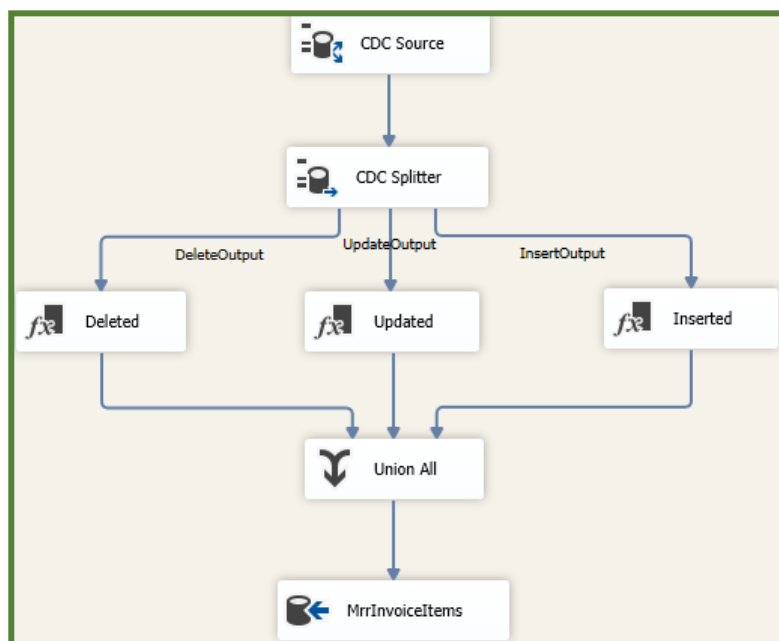
לצורך יצירת טבלת ה-FACT נבצע JOIN בין טבלת INVOICES לטבלת INVOICEITEMS.



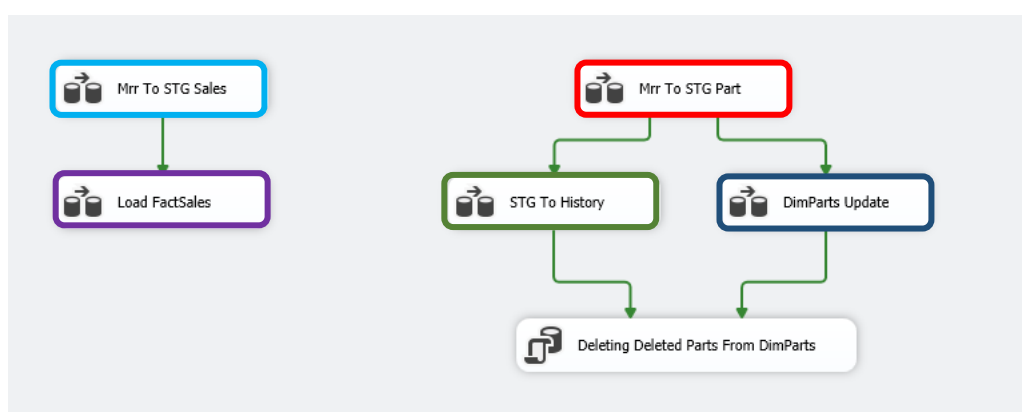
חבילה זו תורץ פעם אחת בלבד לשם יצירת הטבלאות עצמן, המבנה שלהן והמידע הראשוני המגיע מבסיס הנתונים.

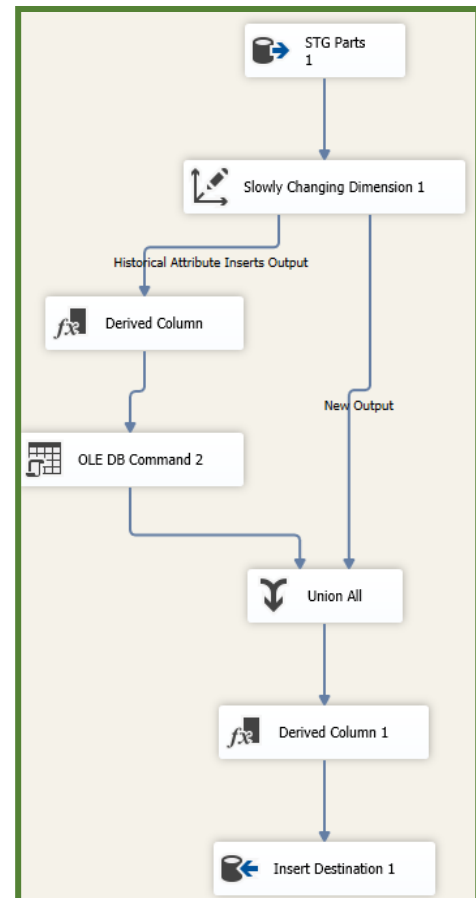
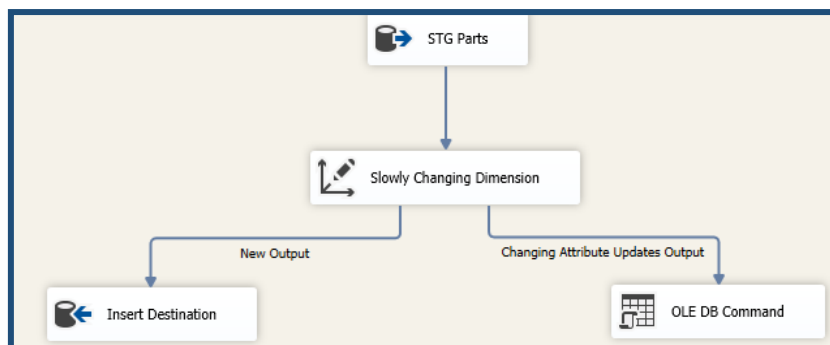
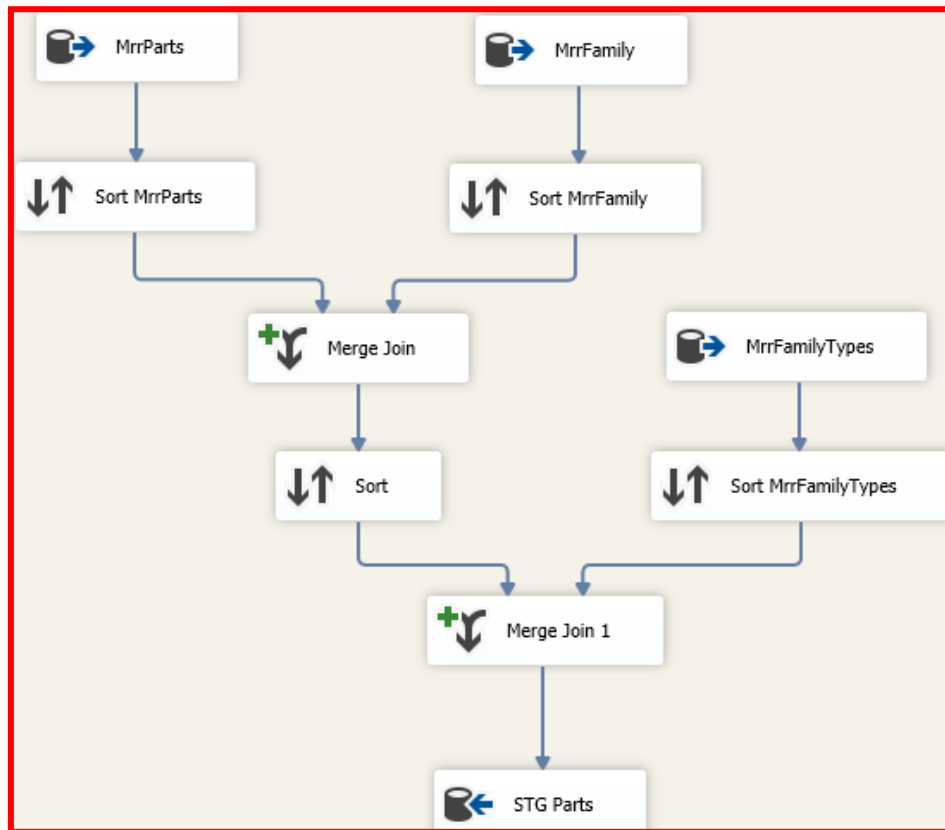
- חבילת ה-Extract בה תחילה נבצע ניקוי לטבלאות ה-mrr- וה-stg הזמניות שלנו (אותן ניצור בהמשך).  
לאחר מכן באמצעות תהליכי CDC ניצור טבלה בבסיס הנתונים התפעולי (cdc\_states) המתעדת את השינויים שנעשו בטבלאות INVOICES ו-INVOICEITEMS מההרצה הקודמת ונכניס את הנתונים לטבלאות ה-mrr המתאימות.  
לאחר יצירת טבלאות ה-mrr נבצע נבצע הפנייה לחבילת ה-Load.



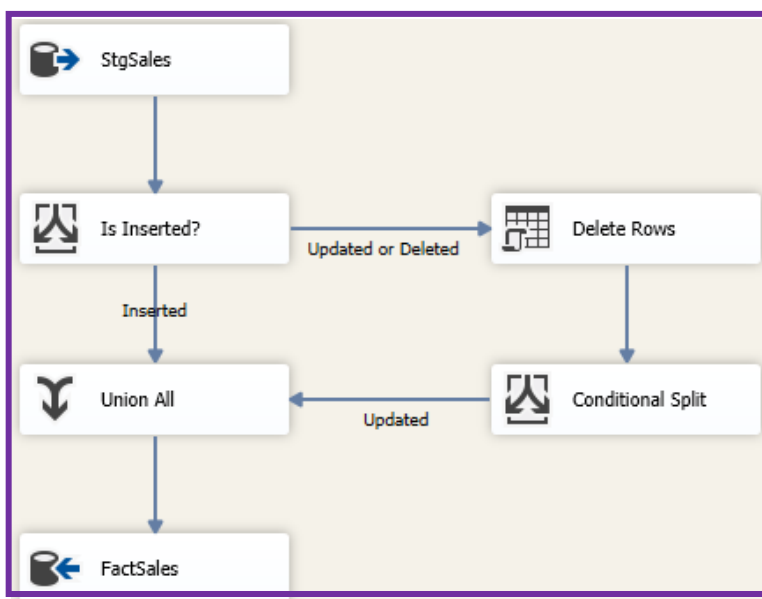
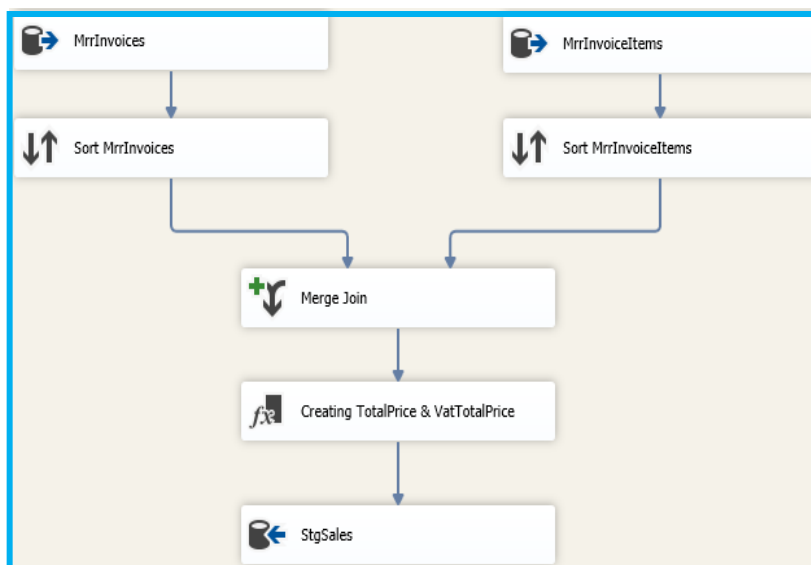


- חבילת ה-Load תבצע שני תהליכים שירוצו בצורה מקבילית.  
**בתהליך הראשון** בו נקח את טבלאות ה-mrr של FAMILY, FAMILYTYPES ו-PART נבצע JOIN ביניהם כדי ליצור טבלת STAGE ולאחר מכן ביצוע הכנסת הנתונים לטבלת DimParts תוך כדי הפעלת תהליך Slowly Changing Dimension מסוג 4 בכדי לעדכן שורות ב-DimParts תוך כדי שמירת ההסטוריה העדכונים לטבלת DimPartsHistory.  
 לאחר שמירת ההסטוריה ועדכון DimParts יבוצע תהליך של מחיקת שורות מ-DimParts במידה ובוצעה מחיקת הבסיס הנתונים התפעולי, המחיקה תתבצע על ידי השוואת DimParts אל StgParts בכדי למצוא שורות הקיימות בטבלת ה-Dim אך לא ב-STG ולמחוק אותן.



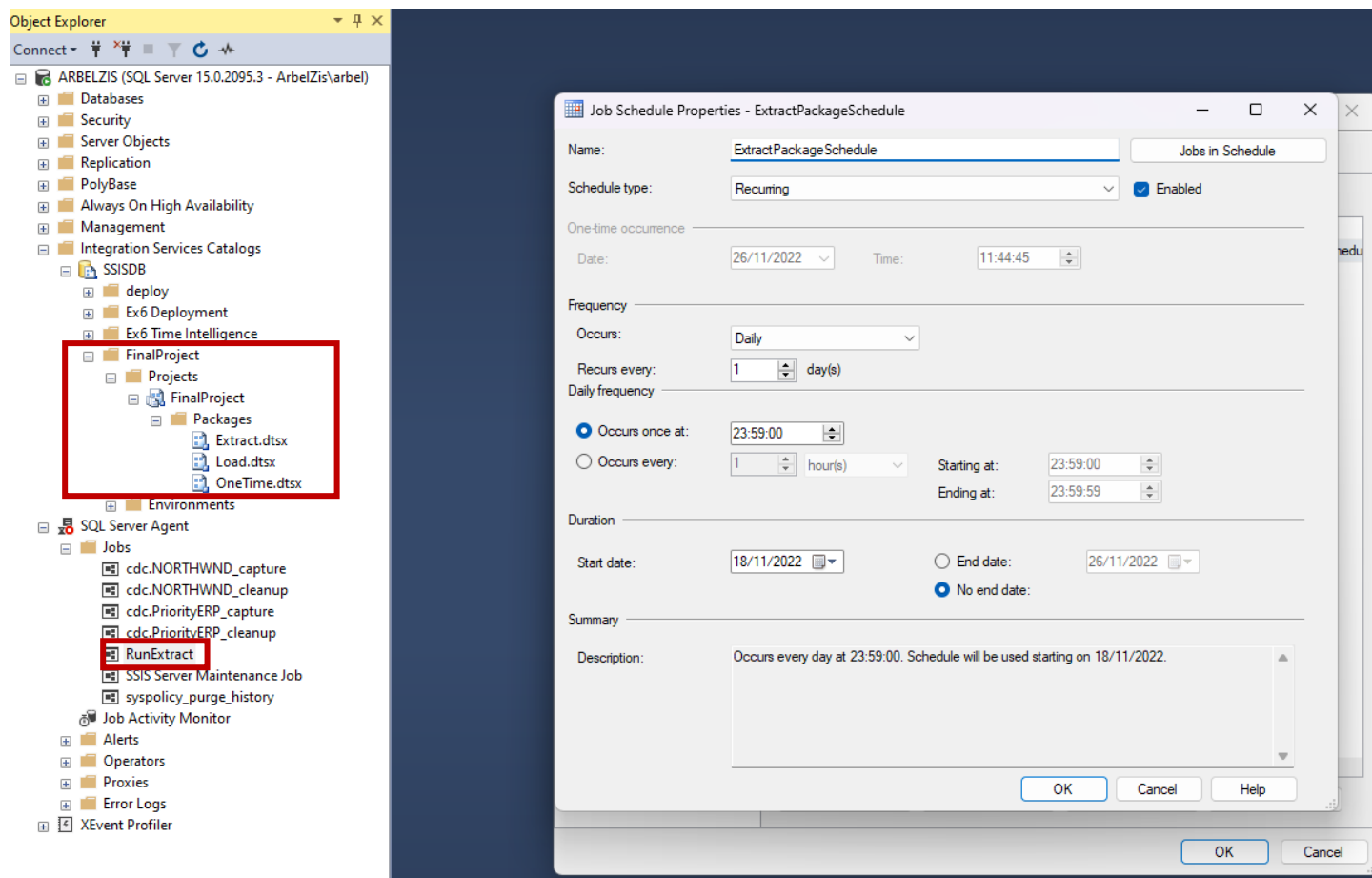


**בתהליך השני** תחילה יבוצע שילוב של טבלאות ה-mrr של INVOICES ו-INVOICEITEMS על ידי JOIN בכדי ליצור טבלת stg ואז העברה אל טבלת ה-fact והחלטה על דרך הפעולה שבה יש לעדכן את הנתונים לפי השינויים שתועדו על ידי ה-CDC (עדכון, הוספה, מחיקה).



לאחר סיום חבילת ה-Load תבוצע חזרה לחבילת ה-Extract שם ה-CDC יבצע סימון של טווח הזמן החדש ממנו יש צורך להתחיל לתעד שינויים חדשים.

לאחר יצירת תהליך ה-ETL בוצע Deploy מ-SSIS אל ה-SSMS לצורך הגדרת גיבוי יומי דרך ה-SQL Agent שתפקידו להריץ את חבילת ה-Extract בכדי לבצע עדכון נתונים על בסיס יומי שאותו נגדיר לפי צרכי הארגון.



### 3.1.3 – תיאור טבלאות הנתונים ב-Datwarehouse

**DimParts** – טבלת מימד המוצרים.

**DimPartsHistory** – טבלת מימד היסטורית המוצרים, טבלה השומרת את כל היסטורית השינויים עבור כל מוצר עם תאריך התחלה ותאריך סיום במידה והשינוי שנעשה כבר לא עדכני.

**DimAgents** – טבלת מימד העובדים.

**DimCustomers** – טבלת מימד הלקוחות.

**FactSales** – טבלת המכירות המורכבת מטבלת ההזמנות (Invoices) וטבלת פירוט ההזמנות (InvoiceItems).

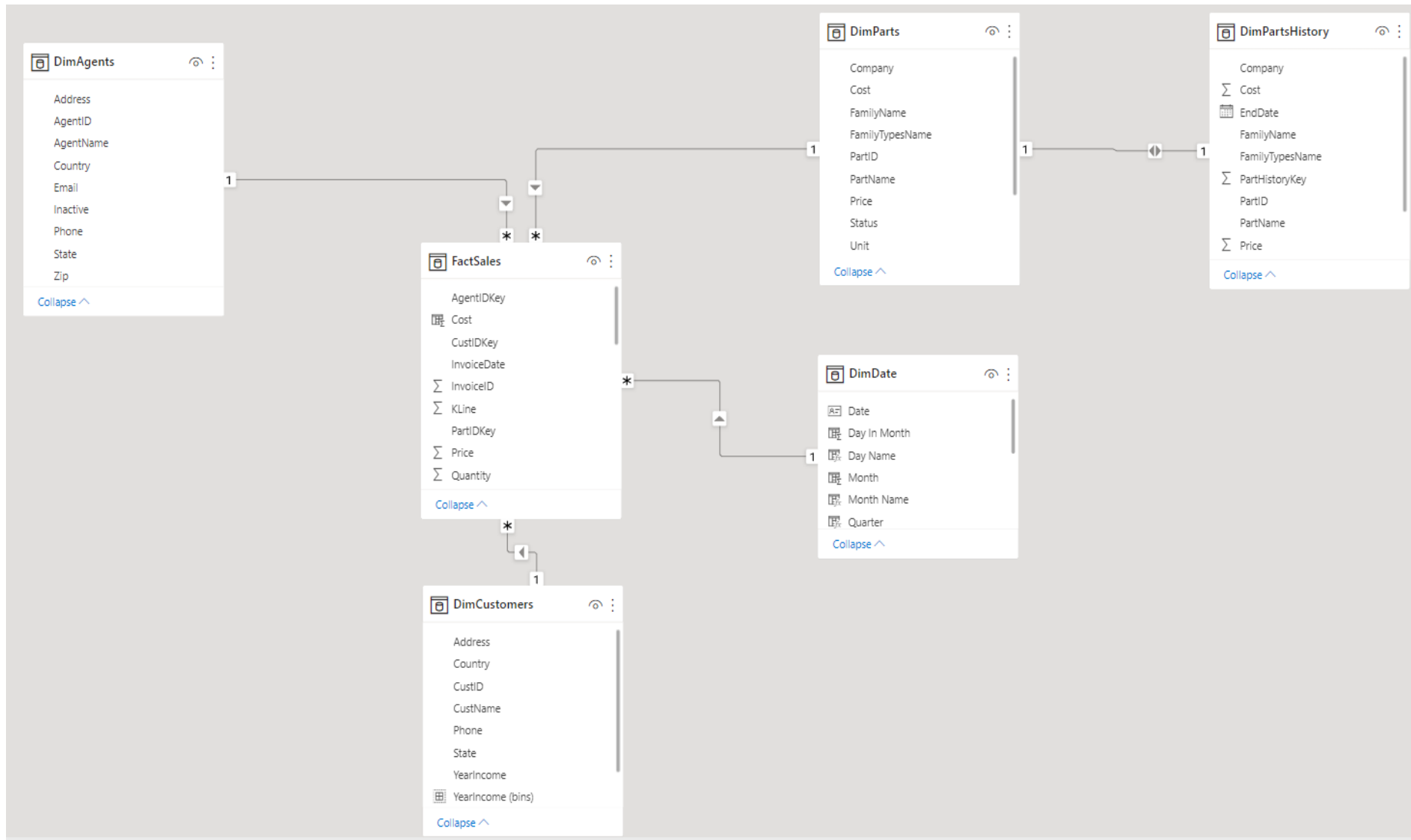
### 3.1.4 – תיאור המדדים, היררכיות, תרשימים והטבלאות

כשלב מקדים ליצירת הדו"ח בוצע ייבוא של הטבלאות משרת ה-SQL ויצירת וחיבור המודל.

לצורך כך נבנתה טבלה חדשה עבור מימד הזמן DimDate ולבסוף חיבור שלה למודל עצמו.

✕ ✓ 1 DimDate = CALENDARAUTO()

Date	Year	Month	Month Name	Quarter	Weekday	Day Name	Day In Month	Week In Year
01/01/2017	2017	1	January	Q1	1	Sunday	1	1
02/01/2017	2017	1	January	Q1	2	Monday	2	1
03/01/2017	2017	1	January	Q1	3	Tuesday	3	1
04/01/2017	2017	1	January	Q1	4	Wednesday	4	1
05/01/2017	2017	1	January	Q1	5	Thursday	5	1
06/01/2017	2017	1	January	Q1	6	Friday	6	1
07/01/2017	2017	1	January	Q1	7	Saturday	7	1
08/01/2017	2017	1	January	Q1	1	Sunday	8	2
09/01/2017	2017	1	January	Q1	2	Monday	9	2
10/01/2017	2017	1	January	Q1	3	Tuesday	10	2
11/01/2017	2017	1	January	Q1	4	Wednesday	11	2
12/01/2017	2017	1	January	Q1	5	Thursday	12	2
13/01/2017	2017	1	January	Q1	6	Friday	13	2
14/01/2017	2017	1	January	Q1	7	Saturday	14	2
15/01/2017	2017	1	January	Q1	1	Sunday	15	3
16/01/2017	2017	1	January	Q1	2	Monday	16	3
17/01/2017	2017	1	January	Q1	3	Tuesday	17	3
18/01/2017	2017	1	January	Q1	4	Wednesday	18	3
19/01/2017	2017	1	January	Q1	5	Thursday	19	3
20/01/2017	2017	1	January	Q1	6	Friday	20	3
21/01/2017	2017	1	January	Q1	7	Saturday	21	3
22/01/2017	2017	1	January	Q1	1	Sunday	22	4





## מדדים:

**Total Sales** -  $SUM(FactSales[TotalPrice])$

**Total Units** -  $SUM(FactSales[Quantity])$

**Total Cost** -  $SUM(FactSales[Cost])$

**Profit Margin** -  $[Total\ Sales] - [Total\ Cost]$

**YTD Sales** -  $TOTALYTD([Total\ Sales], DimDate[Date])$

**LY YTD Sales** -  $CALCULATE([YTD\ Sales], SAMEPERIODLASTYEAR(DimDate[Date]))$

**Sales Diff From Last Year** -  $[Total\ Sales] - [LY\ Sales]$

**Sales Diff From Last Year %** -  $DIVIDE(FactSales[Sales\ Diff\ From\ Last\ Year], [LY\ Sales])$

**MTD Sales** -  $TOTALMTD([Total\ Sales], DimDate[Date])$

**LM MTD** -  $CALCULATE([MTD\ Sales], DATEADD(DimDate[Date], -1, month))$

**MTD Change** -  $DIVIDE([MTD\ Sales] - [LM\ MTD], [LM\ MTD])$

## היררכיות:

### **Time Hierarchy:**

Year

Quarter

Month

Day In Month

### **Product's Category Hierarchy:**

FamilyTypesName

FamilyName

PartName

## גרפים וטבלאות:

### עמוד מכירות

KPI :

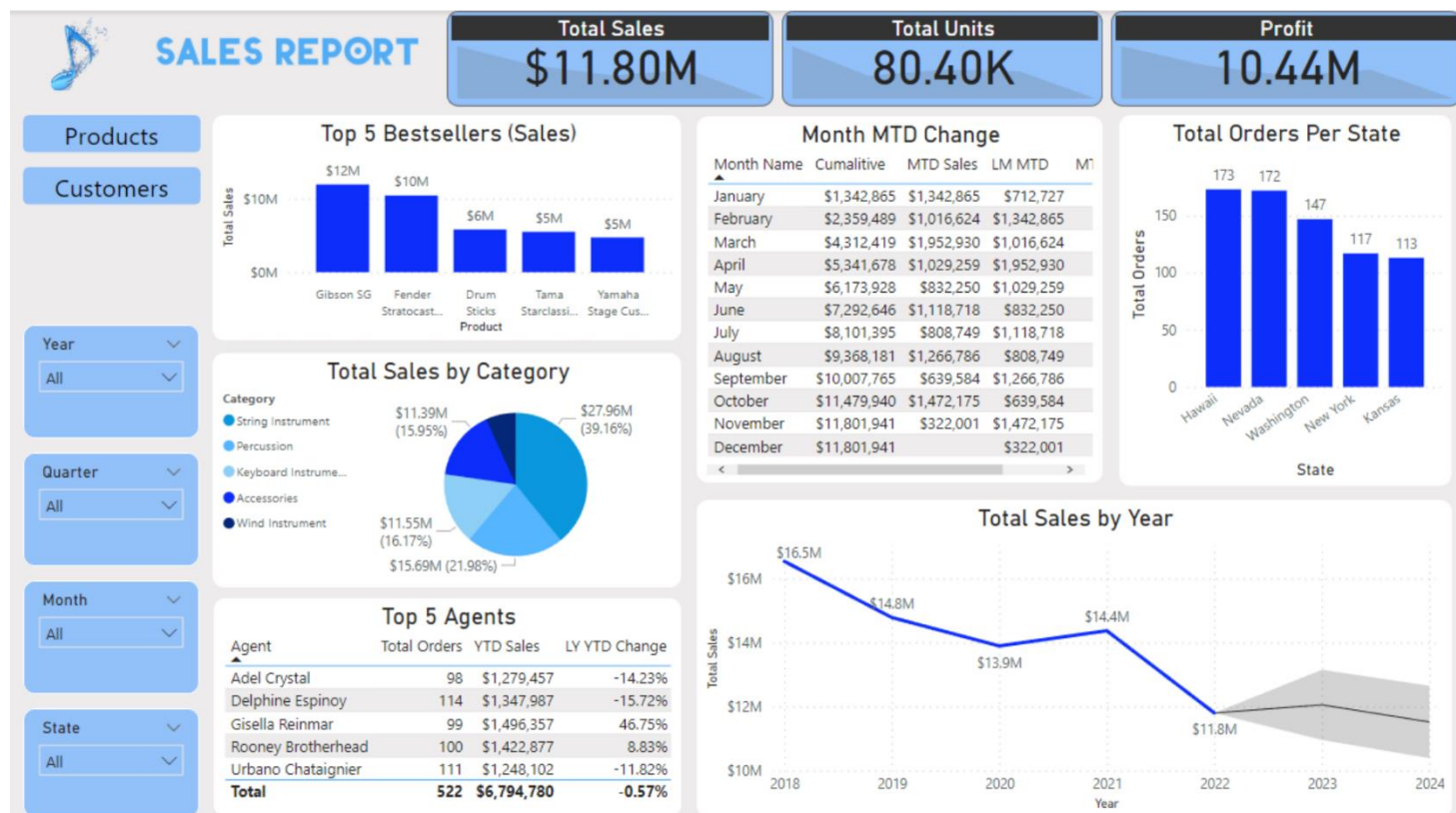
- סך המכירות שנעשו.
- סך היחידות שנמכרו.
- רווח נטו שנעשה.

גרפים :

- 5 המוצרים שהניבו את המכירות הגבוהות ביותר.
- 5 ה-States שבהם נעשו הכי הרבה הזמנות.
- פילוח הקטגוריות עם Drill Down בהיררכיית המוצרים לפי מכירות.
- Line Chart המתאר את המכירות עבור כל שנה עם תחזית לשנתיים הבאות.

טבלאות :

- 5 הסוכנים המוצלחים ביותר לפי מדד YTD Sales עם פירוט כמות הזמנות, מכירות עבור שנה נוכחית והשינוי בתפוקה מהשנה הקודמת לה.
- פירוט המכירות בחודשי השנה ועם פירוט של סכום מצטבר, מכירות לאותו חודש, מכירות לחודש הקודם לו, ואחוז השינוי בין החודש לחודש הקודם לו.



## עמוד לקוחות

גרפים:

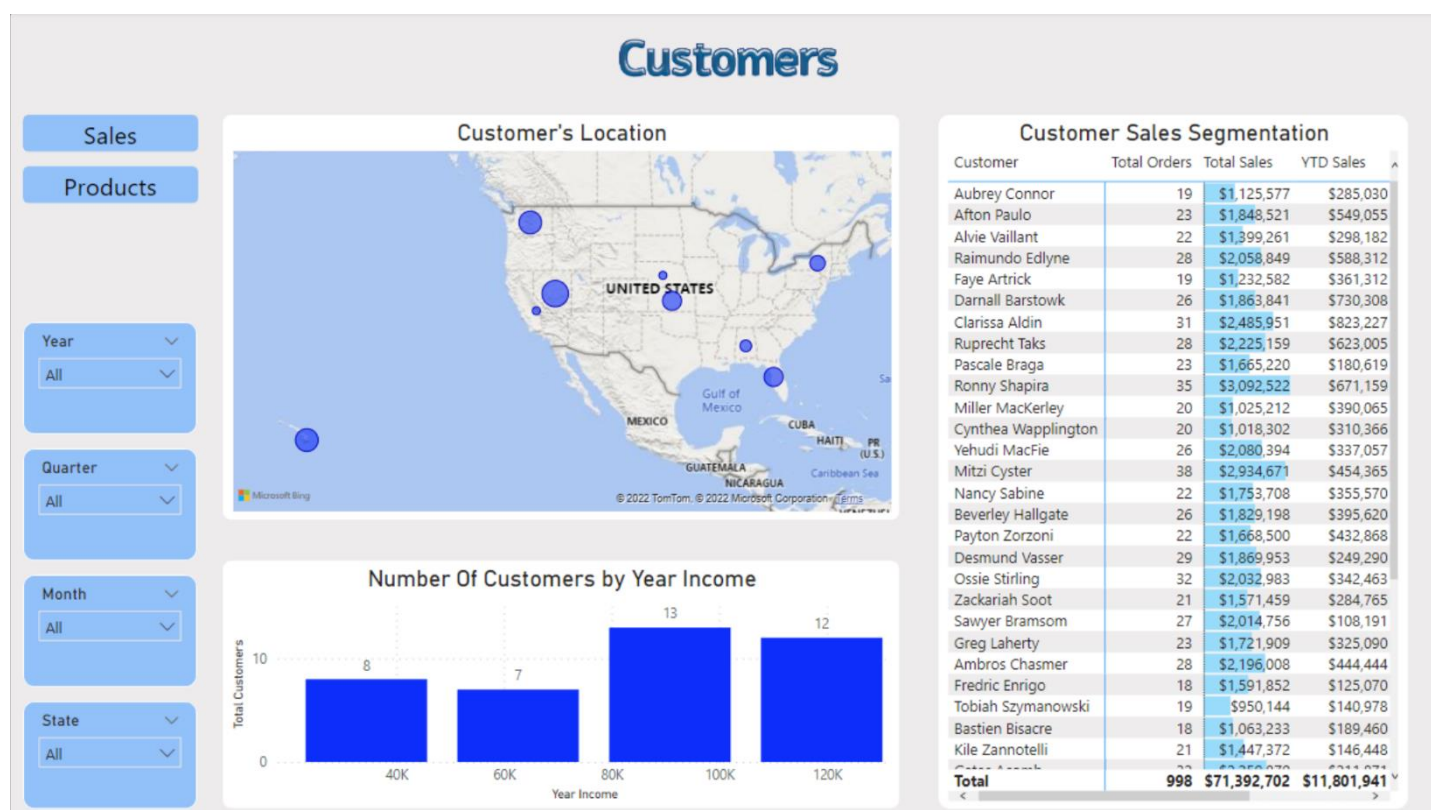
- כמות הלקוחות עבור כל חלוקה שנוצרה להכנסה שנתית של הלקוחות.

מפה:

- כמות הלקוחות עבור כל State.

טבלאות:

- תיאור נתוני לקוחות עם פירוט שם לקוח, כמות הזמנות שביצע, כמות הכסף שבה רכש, רכישות עבור שנה נוכחית ושינוי באחוזים מהרכישות בשנה הקודמת.



## עמוד המוצרים

גרפים:

- תיאור הקשר בין כמות ההזמנות שנעשו בהם הופיע מוצר מסוים לכמות היחידות שנמכרו ממנו. עם שימוש בפונקציית Play Axis של המראה את השינוי בין שנה לשנה.

מפה:

- מפה המראה כמות יחידות שנמכרו עבור כל State עם Tooltip המראה גרף נוסף של פילוח 3 המוצרים עם הכי הרבה יחידות שנמכרו עבור ה-State הנבחר.



\*עבור כל העמודים בדו"ח קיימים סלייסרים של שנה, רבעון, חודש וה-State.

Year
All

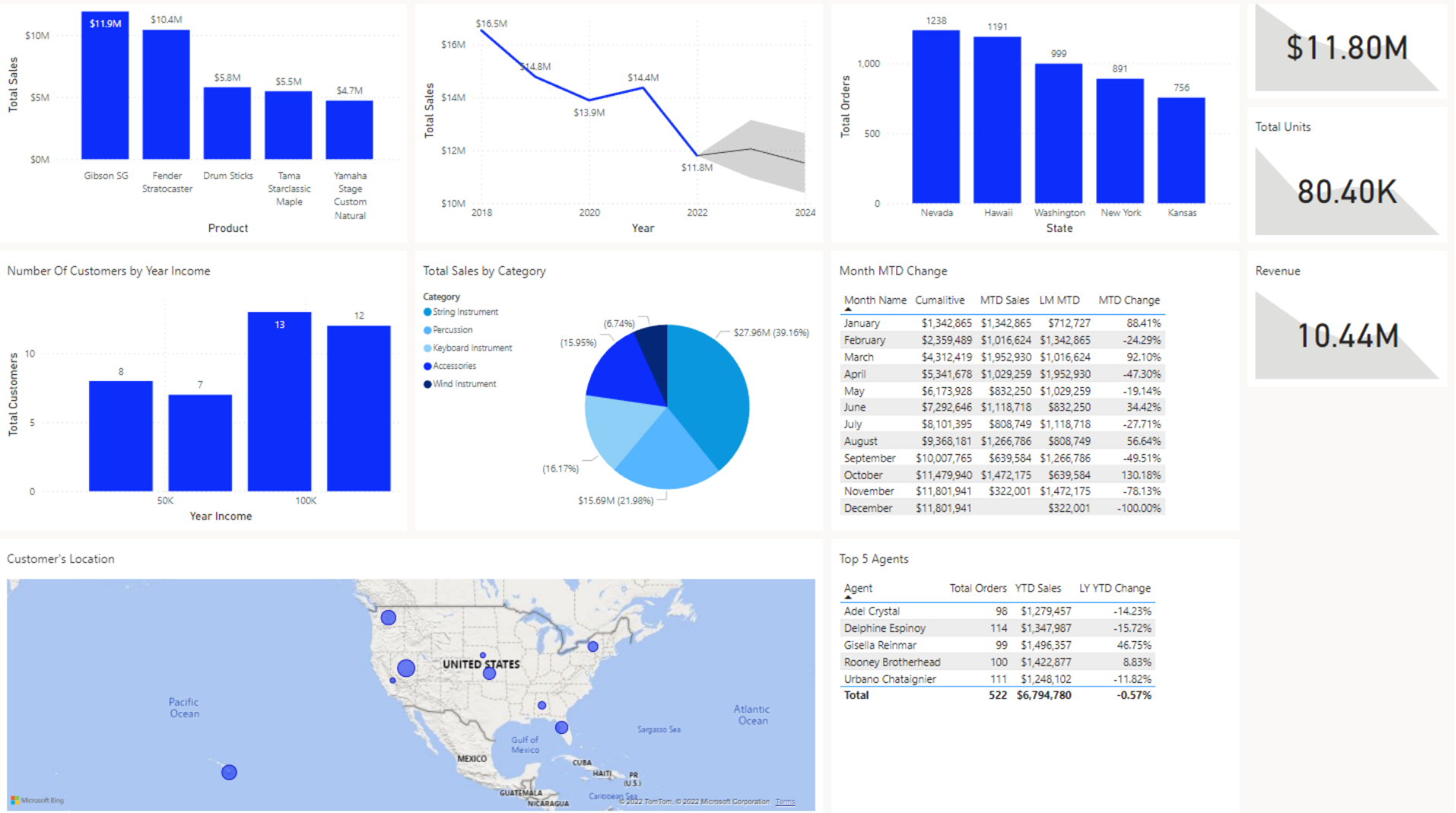
Quarter
All

Month
All

State
All

לאחר יצירת הדוחות בוצע תהליך Publish עבור ה-Power BI Service שם נוצר Dashboard.

**:Dashboard**






לאחר יצירת ה-Dashboard נוצר App עבור אנשי הארגון אליו הוספו ה-Dashboard והדוחות.  
בנוסף, באמצעות ה-Gateway שיצרנו בוצע Schedule refresh על בסיס יומי בשעה 6 בבוקר.

#### Gateway connection

To use a data gateway, make sure the computer is online and the data source is added in [Manage Gateways](#). If you're using an On-premises data gateway (standard mode), please select the corresponding data sources and then click apply.

Use an On-premises or VNet data gateway

☐ On

Gateway	Department	Contact information	Status	Actions
 Personal Gateway			 Running on ARBELZIS	

Apply

Discard

#### Data source credentials

#### Parameters

#### Scheduled refresh

Keep your data up to date

Configure a data refresh schedule to import data from the data source into the dataset. [Learn more](#)

☒ On

Refresh frequency

Daily

Time zone

(UTC+02:00) Jerusalem

Time

6 00 AM

[Add another time](#)

Send refresh failure notifications to

☒ Dataset owner

☐ These contacts:

## 4. תוכנית עבודה

13.11.2022 - תאריך תחילת פרויקט

13.11.2022 – 20.11.2022 - פיתוח תהליך ה-ETL באמצעות Visual Studio SSIS

21.11.2022 – 23.11.2022 - ייצור הדוחות וה-Dashboard באמצעות Power BI

24.11.2022 – 25.11.2022 - סיום כתיבת מסמך האפיון.

25.11.2022 – סיום פרויקט.