

מסמך איפיון – חברת ONLY CLOTHING



תיאור החברה:

חברת ONLY CLOTHING הינה חברה של חנויות בגדים הפרוסה ברחבי ארה"ב. ישנם כ-50 חנויות, וכ-1000 עובדים. החנויות כרגע הן חנויות פיזיות, וייתכן כי תהיה התפתחות בעתיד לכיוון המכירות באונליין.

עיקר המכירות של החברה הם בגדי נשים, גברים וילדים.

תיאור המערכת:

המערכת תציג חישובי הכנסות מול הוצאות של החברה ONLY CLOTHING. בנוסף, המערכת מעניקה ניתוח ברמה גיאוגרפית ובנוסף ניתוח על פני צירי זמן שונים, קטגוריות, מוצרים ומוכרים.

מטרת המערכת:

ניתוח המכירות של החברה והפקת תובנות עסקיות לגבי האופן בו ניתן לשפר את המודל העסקי של החברה באמצעות מדדים שונים.



יעדי המערכת:

בניית DWH שמטרתו לבצע ניתוחים אנליטיים מורכבים על מקורות הנתונים המצויים במערכת התפעולית של PRIORITY ERP. בניית ה-DWH כוללת יצירת טבלאות DIMENSION וטבלת FACT אחת. בנוסף, יצירת דוחות רלוונטים עם חיתוכים שונים לגבי נתוני החברה.

פירוט הטבלאות DWH:

- DimAgents: טבלה השומרת את פרטי כל הסוכנים בחברה.
- DimDate: טבלת זמן.
- DimCategories: טבלה השומרת את כל פרטי קטגוריות האב השונות של מוצרי החברה.
- DimSubCategories: טבלה השומרת את כל פרטי תת הקטגוריות (קטגוריות הבנים) השונות של מוצרי החברה.
- DimCountries: טבלה השומרת את פרטי המדינות השונות.
- DimCustomers: טבלה השומרת את כל פרטי הלקוחות של החברה.
- DimProducts: טבלה השומרת את כל פרטי המוצרים השונים שקיימים בחברה.
- DimStores: טבלה השומרת את כל פרטי החנויות השונות של החברה.
- FactSales: טבלה המאגדת בתוכה את כל הגורמים הקשורים במכירות ובפרטי המכירות שהתבצעו.

טכנולוגיות בשימוש:

- מסד הנתונים שייך לחברת Priority ונמצא על שרת MSSQL.
- מימוש תהליך ה-ETL ע"י תוכנת ה-SSIS. בתהליך זה תתבצע טעינה אינקרימנטלית תוך שימוש בטבלאות MRR, טבלאות STG ולבסוף טעינה אל טבלאות היעד שבתוך ה-DWH, הנמצאות על שרת MSSQL.
- ניתוחים אנליטיים והפקת דוחות ע"י תוכנת POWER BI על ה-DWH של החברה. יצירת מדדים בשפת DAX על מנת להפיק חישובים ותובנות על נתוני החברה.

מקורות נתונים:

- רוב הטבלאות מגיעות ממסד הנתונים ERP של חברת Priority שנמצא על שרת MSSQL.
- טבלת DimDate נוצרה ע"י סקריפט המורץ על שרת MSSQL.
- יצירת טבלאות CDC ע"י הרצת פרוצדורה על מסד הנתונים ERD של חברת Priority וכן הרצת פרוצדורה נוספת על כל הטבלאות שעליהן נרצה לבצע את הטעינה האינקרימנטלית.



תהליך ה-ETL:

התהליך מתחלק לארבעה חלקים, כאשר מטרת החלק הראשון היא טעינה היסטורית של כל הנתונים המצויים במערכת התפעולית וטעינתם אל ה-DWH.

שאר החלקים מתארים את תהליך ה-ETL. התהליך מתחיל לאחר סימון Mark CDC Start, אחריו תתחיל הטעינה האינקרימנטלית.

1. חילוץ הנתונים מטבלאות המקור שבמסד הנתונים PRIORITYERP. חילוץ הנתונים נעשה באופן אינקרימנטלי, כאשר טבלאות המקור אלה טבלאות ה-CDC שנוצרו עבור כל טבלה. באותה טעינה, נרצה לראות את השינויים נטענים אל טבלאות ה-MRR. שתפקידן להוות שיקוף של השינויים באותה טעינה בסוף שלב החילוץ נסמן את הטווח הזמן החדש של הטעינה האינקרימנטלית.
2. לאחר מכן, נבצע טרנספורמציות על הנתונים המגיעים מתוך טבלאות ה-MRR.
 - נעדכן את טיפוסים הנתונים של השדות
 - נעדכן את שמות השדות אל השמות הסופיים שבטבלאות שב-DWH
 - נמזג את הטבלאות הבאות: INVOICES_MRR יחד עם INVOICEITEMS_MRR וכן נאפיין את סוגי פקודות ה-DML השונות שקרו בחלון הזמן של טעינה זו. המטרה היא לבצע הכנה של הנתונים לקראת הטעינה הסופית אל טבלת ה-FACT ב-DWH.בשלב זה מקור הנתונים שלנו הן טבלאות ה-MRR והיעד הן טבלאות ה-STG.
3. בשלב האחרון נבצע טעינה אל טבלאות היעד ב-DWH, אלו הן טבלאות ה-DIM וטבלת ה-FACT.

בשלב זה תעשה בדיקה והענקת מענה מתאים לפי פקודות ה-DML השונות שתועדו בטבלאות ה-CDC.

מלבד השימוש בטבלאות ה-CDC, נעשה גם שימוש בסוג נוסף של טעינה אינקרימנטלית – שימוש ב-SCD על טבלת DimAgents, הסוג שנבחר הוא SCD TYPE 2.

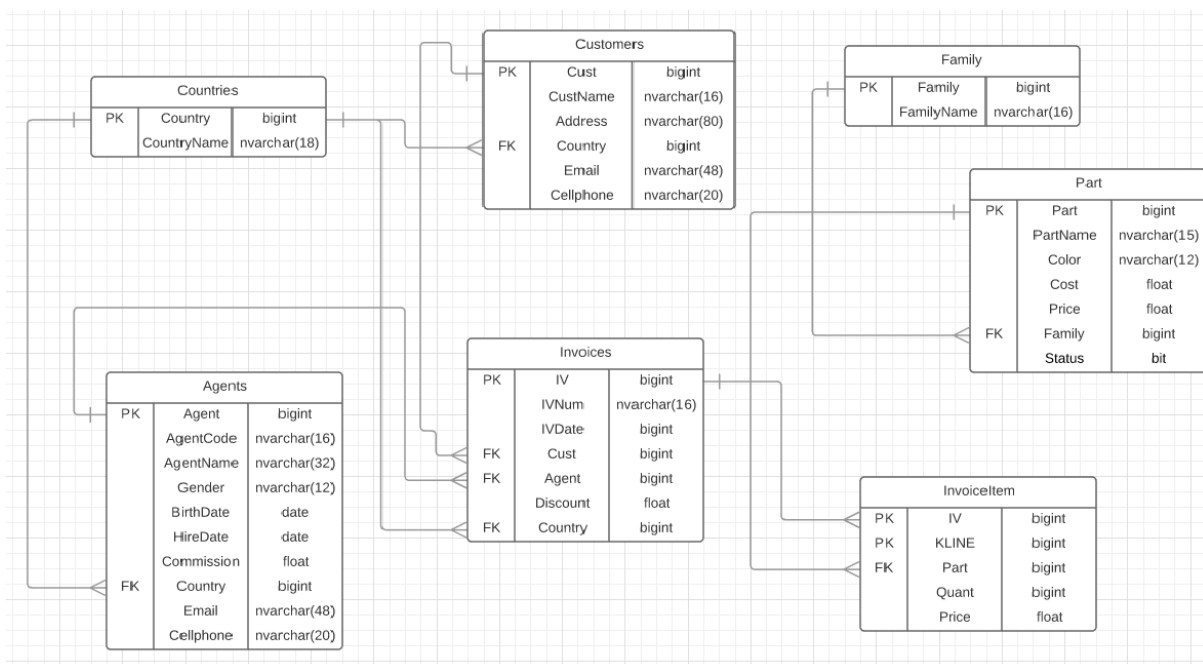
לצורך כך הוספנו לטבלת DimAgents שני שדות נוספים: StartDate ו-EndDate. SCD TYPE 2 מטפל בשמירה היסטורית, ז"א כאשר מתבצעת פקודת DML מסוג UPDATE או DELETE אז ישנו תיעוד של השינוי באמצעות שמירת תאריך הסיום שבו התרחש השינוי על הרשומה. כברירת מחדל תאריך הסיום הוא NULL ובאופן זה מעיד על כך כי הרשומה הנ"ל היא הרשומה הרלוונטית ביותר (כל עוד לא התרחשו על הרשומה שינויים) הערך של StartDate התקבל ממערכת ה-SSIS.

תזמון המערכת:

תהליך ה-ETL מתוזמן לרוץ בכל יום בשעות הלילה (00:00) תוך רצון להעניק עבודה רציפה ללא הפרעות למשתמשי הקצה השונים. תהליך זה הוגדר כ-JOB ב-SQL AGENT שתפקידו לבצע את התהליך בקפיצות הזמן שהוגדרו מראש.



סכמת ERD של מסד הנתונים PriorityERP:





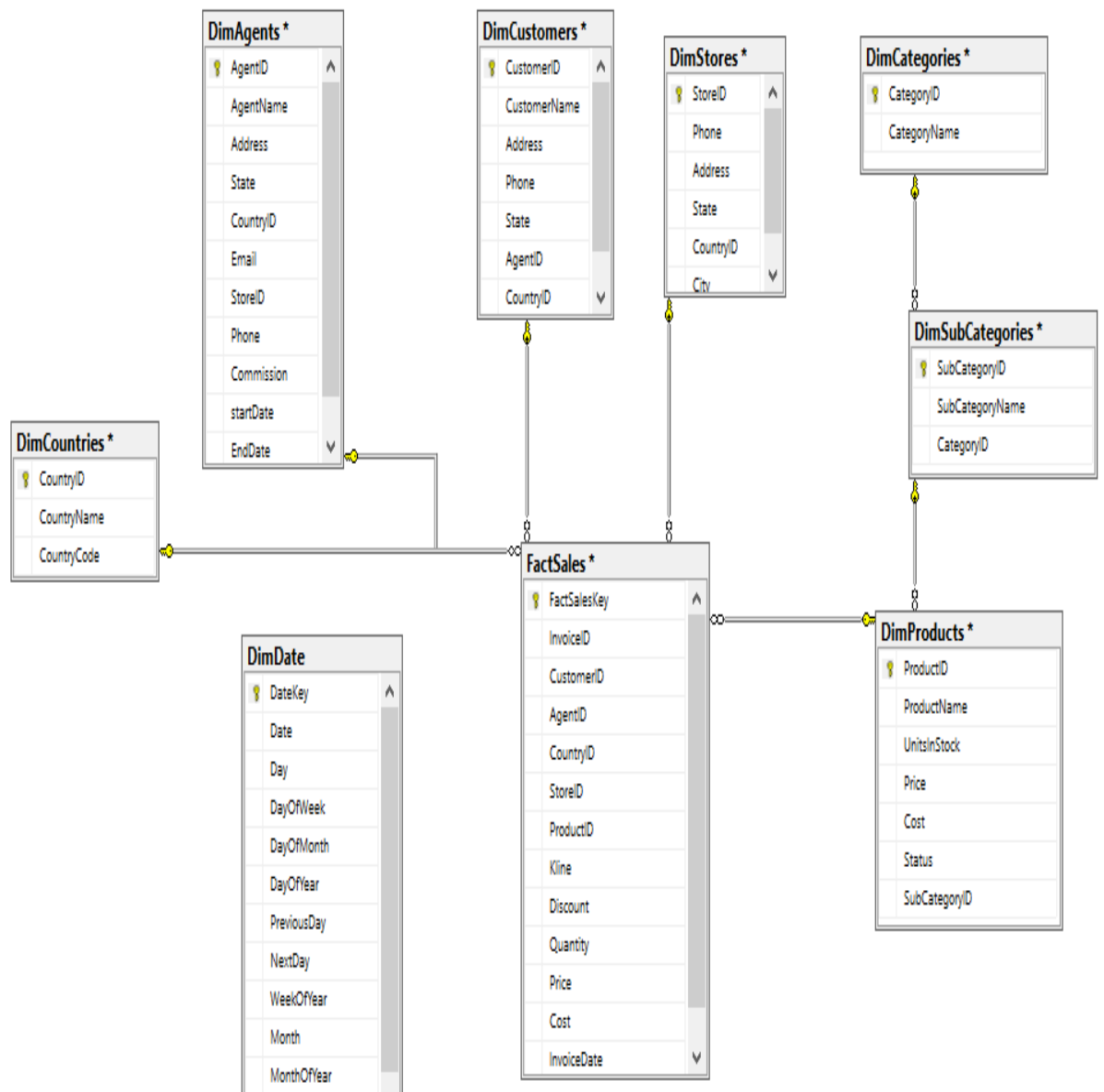
_:Source To Target

Database	MRR	STG	DWH
AGENTS.ADDRESS		STG_DimAgents.Address	DimAgents.Address
AGENTS.AGENT		STG_DimAgents.AgentID	DimAgents.AgentID
AGENTS.AGENTNAME		STG_DimAgents.AgentName	DimAgents.AgentName
AGENTS.COMMISSION		STG_DimAgents.Commission	DimAgents.Commission
AGENTS.EMAIL		STG_DimAgents.Email	DimAgents.Email
AGENTS.PHONE		STG_DimAgents.Phone	DimAgents.Phone
AGENTS.COUNTRY		STG_DimAgents.CountryID	DimAgents.CountryID
AGENTS.STATE		STG_DimAgents.State	DimAgents.State
AGENT.BRANCH		STG_DimAgents.StoreID	DimAgents.StoreID
			DimAgents.StartDate
			DimAgents.EndDate
			FactSales.CountryID
			FactSales.StoreID
BRANCHES.BRANCH	MRR_ BRANCHES.BRANCH	STG_DimStores.StoreID	DimStores.StoreID
BRANCHES.PHONE	MRR_ BRANCHES.PHONE	STG_DimStores.Phone	DimStores.Phone
BRACHES.ADDRESS	MRR_ BRACHES.ADDRESS	STG_DimStores.Address	DimStores.Address
BRANCHES.STATE	MRR_ BRANCHES.STATE	STG_DimStores.State	DimStores.State
BRANCHES.COUNTRY	MRR_ BRANCHES.COUNTRY	STG_DimStores.CountryID	DimStores.CountryID
BRANCHES.CITY	MRR_ BRANCHES.CITY	STG_DimStores.City	DimStores.City
COUNTRIES.COUNTRY	MRR_ COUNTRIES.COUNTRY	STG_DimCountries.CountryID	DimCountries.CountryID
COUNTRY.COUNTRYNAME	MRR_ COUNTRY.COUNTRYNAME	STG_DimCountries.CountryName	DimCountries.CountryName
CUSTOMERS.ADDRESS	MRR_CUSTOMERS.ADDRESS	STG_DimCustomers.Address	DimCustomers.Address
CUSTOMERS.AGENT	MRR_CUSTOMERS.AGENT	STG_DimCustomers.AgentID	DimCustomers.AgentID
CUSTOMERS.COUNTRY	MRR_CUSTOMERS.COUNTRY	STG_DimCustomers.CountryID	DimCustomers.CountryID
CUSTOMERS.CUST	MRR_CUSTOMERS.CUST	STG_DimCustomers.CustomerID	DimCustomers.CustomerID
CUSTOMERS.CUSTNAME	MRR_CUSTOMERS.CUSTNAME	STG_DimCustomers.CustomerName	DimCustomers.CustomerName
CUSTOMERS.PHONE	MRR_CUSTOMERS.PHONE	STG_DimCustomers.Phone	DimCustomers.Phone
CUSTOMERS.STATE	MRR_CUSTOMERS.STATE	STG_DimCustomers.State	DimCustomers.State

FAMILY.FAMILY	MRR_FAMILY.FAMILY	STG_DimSubCategory.SubCategoryID	DimSubCategory.SubCategoryID
FAMILY.FAMILYNAME	MRR_FAMILY.FAMILYNAME	STG_DimSubCategory.SubCategoryName	DimSubCategory.SubCategoryName
FAMILY.FAMILYTYPE	MRR_FAMILY.FAMILYTYPE	STG_DimSubCategory.CategoryID	DimSubCategory.CategoryID
FAMILYTYPES.FAMILYTYPE	MRR_FAMILYTYPES.FAMILYTYPE	STG_DimCategory.CategoryID	DimCategory.CategoryID
FAMILYTYPES.FTNAME	MRR_FAMILYTYPES.FTNAME	STG_DimCategory.CategoryName	DimCategory.CategoryName
INVOICEITEMS.COUNTRY	MRR_INVOICEITEMS.COUNTRY	STG_FactSales.CountryID	FactSales.CountryID
INVOICEITEMS.IV	MRR_INVOICEITEMS.IV	STG_FactSales.InvoiceID	FactSales.InvoiceID
INVOICEITEMS.KLINE	MRR_INVOICEITEMS.KLINE	STG_INVOICEITEMS.KLINE	FactSales.Kline
INVOICEITEMS.PART	MRR_INVOICEITEMS.PART	STG_FactSales.ProductID	FactSales.ProductID
INVOICEITEMS.PRICE	MRR_INVOICEITEMS.PRICE	STG_FactSales.Price	FactSales.Price
INVOICEITEMS.QUANT	MRR_INVOICEITEMS.QUANT	STG_FactSales.Quantity	FactSales.Quantity
INVOICES.AGENT	MRR_INVOICES.AGENT	STG_FactSales.AgentID	FactSales.AgentID
INVOICES.CUST	MRR_INVOICES.CUST	STG_FactSales.CustomerID	FactSales.CustomerID
INVOICES.DISCOUNT	MRR_INVOICES.DISCOUNT	STG_FactSales.Discount	FactSales.Discount
INVOICES.IV	MRR_INVOICES.IV	STG_FactSales.InvoiceID	FactSales.InvoiceID
INVOICES.IVDATE	MRR_INVOICES.IVDATE	STG_FactSales.InvoiceDate	FactSales.InvoiceDate
PART.COST	MRR_PART.COST	STG_DimProducts.Cost	DimProducts.Cost
PART.FAMILY	MRR_PART.FAMILY	STG_DimProducts.SubCategoryID	DimProducts.SubCategoryID
PART.PART	MRR_PART.PART	STG_DimProducts.ProductID	DimProducts.ProductID
PART.PARTNAME	MRR_PART.PARTNAME	STG_DimProducts.ProductName	DimProducts.ProductName
PART.PRICE	MRR_PART.PRICE	STG_DimProducts.Price	DimProducts.Price
PART.STATUS	MRR_PART.STATUS	STG_DimProducts.Status	DimProducts.Status
PART.UNIT	MRR_PART.UNIT	STG_DimProducts.UnitsInStock	DimProducts.UnitsInStock



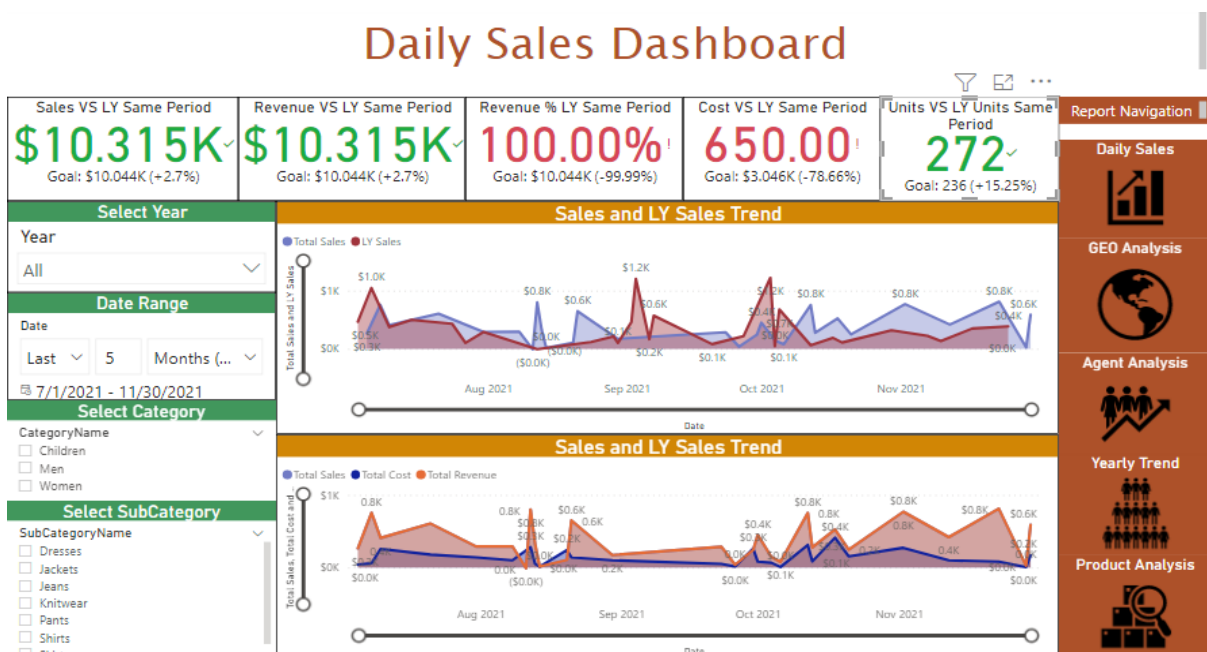
:DWH ERD





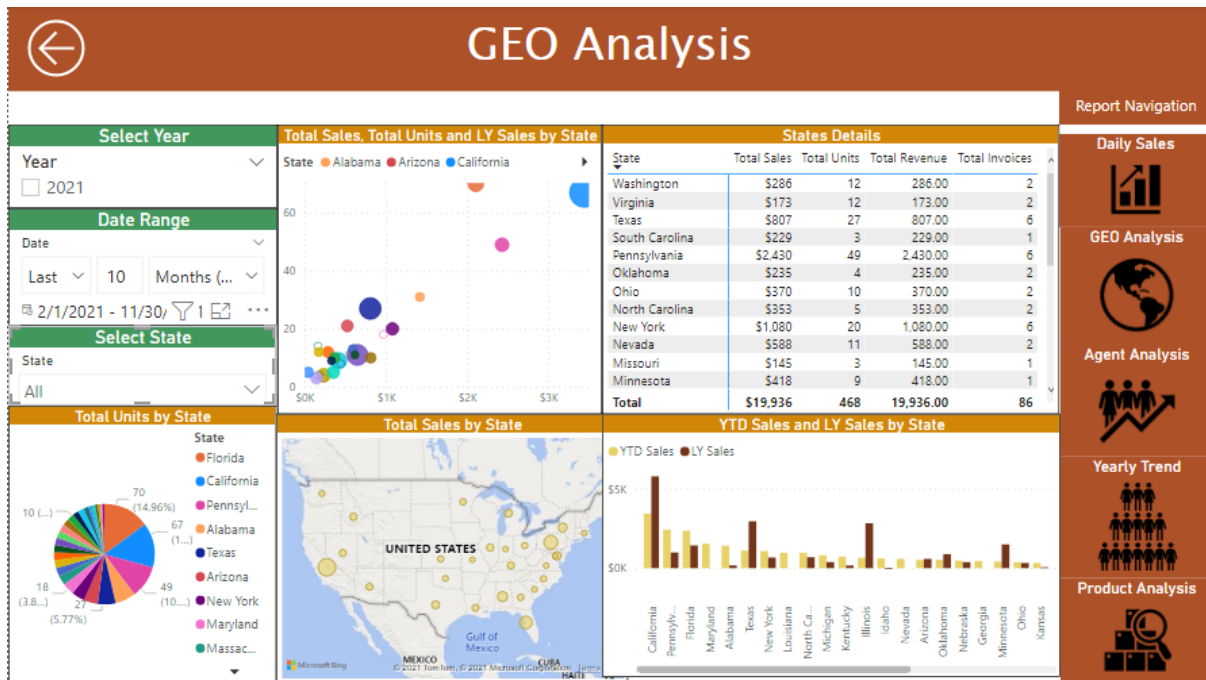
מסכי המערכת :POWER BI

1. **Daily Sales Dashboard**: הלוח מתאר את השבוע אחרון של המכירות, רווחים, כמויות ועלויות. שימוש בKPI : סה"כ מכירות, סה"כ הוצאות, סה"כ רווח, אחוז רווח מסה"כ מכירות, סה"כ עלות וסה"כ כמויות לפי חתך הזמן המוגדר. גרף השוואה בין מכירות בתקופה המוגדרת השנה מול התקופה המוגדרת בשנה שעברה בחיתוך יומי.





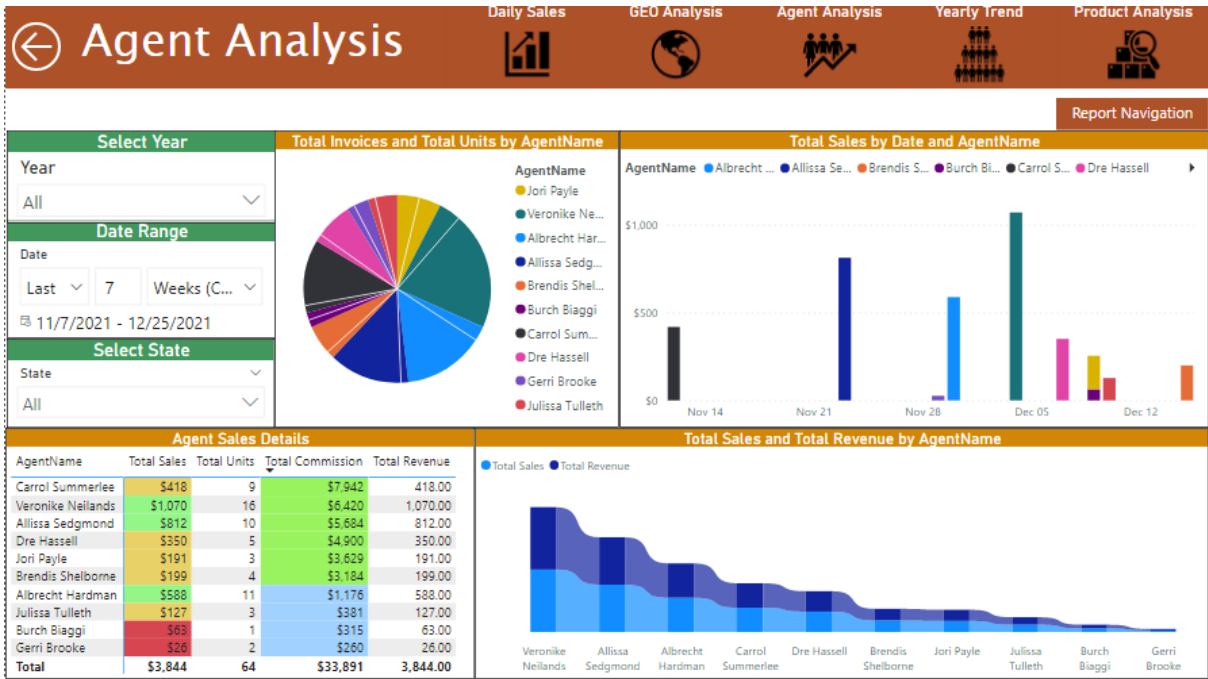
2. **GEO Analysis:** תצוגה של ה States השונים, עם פירוט לגבי כל State ברמה של סה"כ מכירות, סה"כ פריטים שהוזמנו, בה"כ רווח, סה"כ הזמנות. גרף מכירות בחתך של States לעומת שנה שעברה, גרף המציד סה"כ פריטים, סה"כ סכום הזמנות עבור כל State. מפה גיאוגרפית המתארת את ריכוז המכירות בין המדינות השונות בארה"ב. גרף פאי המציג את אחוז הפרטים שנקנו עבור כל State.





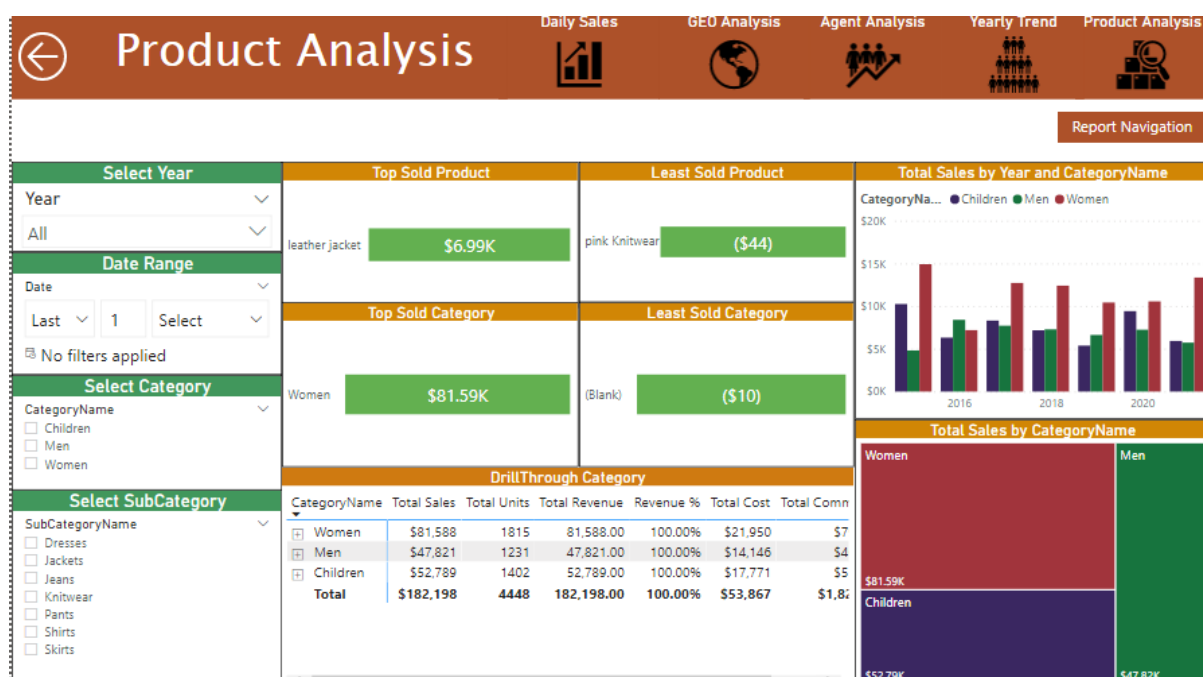
3. **Agent Analysis:** תצוגה של השבוע האחרון, עם אפשרויות סינון. ישנו גרף המציג את סה"כ מכירות וסה"כ רווחים עם אפשרות לבצע drill down עבור ממד הזמן – תצוגה של סה"כ רווח ומכירות לאורך אותו פרק זמן שנקבע בסינון. ישנו גרף פאי המתאר את אחוז ההזמנות שביצע כל עובד ביחד לסה"כ ההזמנות.

ישנה טבלה המציגה את פרטי כל הסוכנים, עם conditional formatting עבור העמודות הבאות: Total Sales – אם סך המכירות הוא בין \$100-\$500 אז הרקע של הערך יצבע בצהוב, אם סך המכירות הוא בין \$500-\$1500 אז הרקע שך הערך יצבע בירוק אם סך המכירות הוא מתחת ל\$100 אז סך המכירות יצבע באדום.

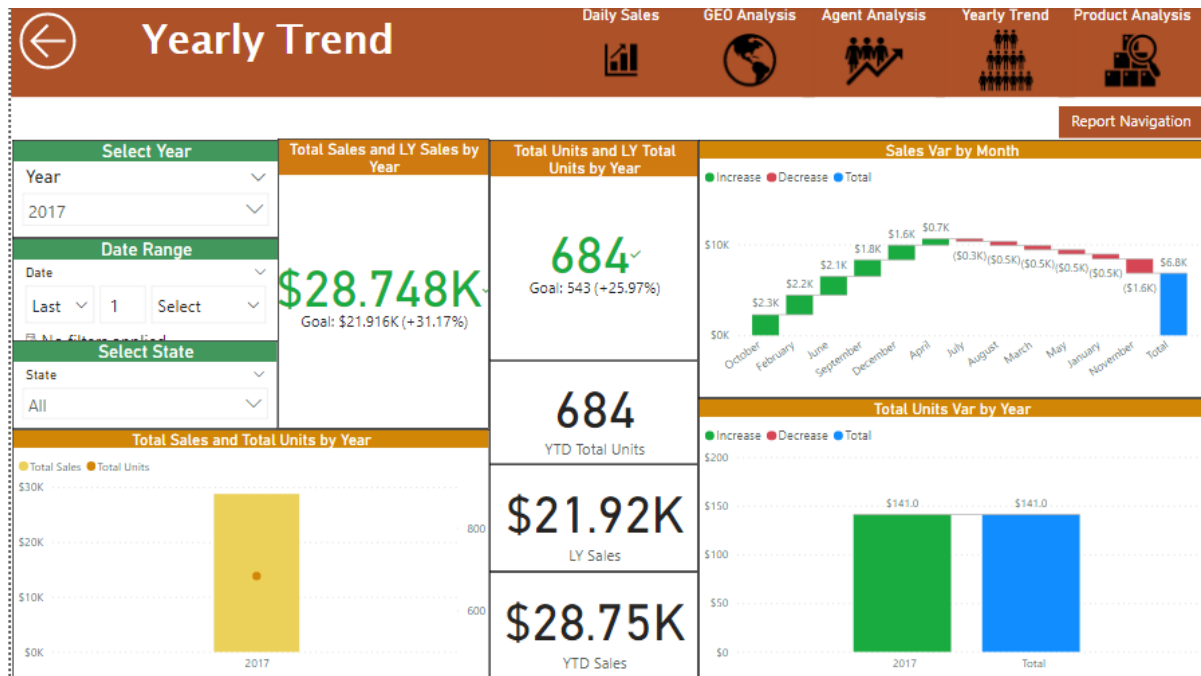




4. **Product Analysis**: תצוגה של הזמן האחרון. רוחים מחולקים לפי קטגוריות אב עם אופציה ל-drill down אל תוך תתי קטגוריות (tree map) בנוסף ישנן אפשרויות סינון ברמת השנה, תאריך, קטגוריית אב וקטגוריית בן. ישנם גרים המציגים את המוצרים הכי מוכרים, את הקטגוריות הכי נמכרות לצד הקטגוריה שהכי פחות נמכרת והמוצר שהכי פחות נמכר. בנוסף ישנה טבלה שדרכה ניתן לעשות מעבר דרך קטגוריות האב, עם פירוט תתי הקטגוריות השייכות לכל קטגוריה. ישנן עמודות המציגות סה"כ רווח, בה"כ מכירות, סה"כ יחידות שנמכרו וסה"כ עמלה עבור כל קטגוריה.



5. **Yearly Trend**: תחזיות שנתיות, עם KPI ומסננים של שנה, תאריך, STATE.



:Measures

YTD Total Units = TOTALYTD([Total Units], 'DimDate'[Date])

YTD Sales = TOTALYTD([Total Sales], 'DimDate'[Date])

Total Commission = SUM(FactSales[Commission])

Total Cost = SUM(FactSales[TotalCost])

Total Discount = SUM(FactSales[DiscountCost])

Total Revenue = SUM(FactSales[Revenue])

Total Sales = SUM(FactSales[SalesAmount])

Total Units = SUM(FactSales[Quantity])

Total Units Var = [Total Units] - [LY Total Units]

Sales Var = [Total Sales] - [LY Sales]

Revenue % = [Total Revenue]/[Total Sales]

LY YTD Total Units = CALCULATE([YTD Total Units], SAMEPERIODLASTYEAR('DimDate'[Date]))

LY YTD Sales = CALCULATE([YTD Sales], SAMEPERIODLASTYEAR('DimDate'[Date]))

LY Total Units = CALCULATE([Total Units], SAMEPERIODLASTYEAR('DimDate'[Date]))

LY Total Sales = CALCULATE([Total Sales], SAMEPERIODLASTYEAR('DimDate'[Date]))

LY Sales = CALCULATE([Total Sales], SAMEPERIODLASTYEAR('DimDate'[Date]))

LY Revenue = CALCULATE([Total Revenue], SAMEPERIODLASTYEAR('DimDate'[Date]))

LY Discount = CALCULATE([Total Discount], SAMEPERIODLASTYEAR('DimDate'[Date]))

LY Cost = CALCULATE([Total Cost], SAMEPERIODLASTYEAR('DimDate'[Date]))

LY Commission = CALCULATE([Total Commission], SAMEPERIODLASTYEAR('DimDate'[Date]))

:Calculated Columns

Commission = RELATED(DimAgents[Commission]) * FactSales[Revenue]

DiscountCost = FactSales[SalesAmount] * FactSales[Discount]

Revenue = FactSales[SalesAmount] - FactSales[DiscountCost] - FactSales[TotalCost]

SalesAmount = (FactSales[Price] * FactSales[Quantity])

$\text{TotalCost} = \text{FactSales}[\text{Cost}] * \text{FactSales}[\text{Quantity}]$