**Fabric AI Hackathon Project**

1. **Problem Statement**
2. **Alignment to Hackathon Category**
3. **Solution Components**
4. **Data Refinement -> Using Excel**

**b) Data Analysis & Exploration -> Use of Co-Pilot for Data Engineering**

1. **Power BI Reports -> User of Co-Pilot for Power BI**
2. **Sending data to SQL Server -> Use of Co-Pilot for Data Factory**
3. **Problem Statement: -**

We will be doing the Economic Impact Analysis

1. We will analyze various Factors of the Economy like unemployment, GDP, cpi, mortgage
2. We will be predicting the mortgage rates for the next 3 quarters.
3. We will also analyze sales of 10 companies during this period.

DataSet:-

We downloaded the Economic Impact Analysis from Kaggle for this Hackathon.

The dataset consists of 3 files:-

1. Economic\_data.csv ->

Its contains data for each quarter ,

and for each quarter it has unemployment\_rate, gdp\_value,cpi,mortgate\_rate

1. Companies.csv

It contains a list of 10 companies

1. Purchase.csv ->

It contains the daily data of product category and revenue for each company

1. Alignment to Hackathon Category :-

**This Project Aligns to Hackathon Category  Best use of Copilot for Microsoft Fabric .**

**In this project we have used**

1. **Co-pilot for Data Engineering**
2. **Co-pilot for Power BI**
3. **Co pilot for Data Factory**

**Pre-requisites for Project :-**

1. **Sign for M365 E5 trial Subscription.**
2. **Create a Microsoft Fabric F1 Capactiy for Co-pilot**
3. **Create a SQL Server database for sending data from Lakehouse to SQL Server using DataFlow gen 2**

A screenshot of a computer

Description automatically generated

1. Solution components
2. Data Refinement

We did Data Refinement in excel

Added columns quarter , year &quaterandyear based on date in purchases and economic data csv.

**User of Fabric Components** :-

Workspace

Create a Fabric Workspace FabricCoPilotWorkspace

LakeHouse

Create a new LakeHouse EconomicDataLakeHouse

Click on Files -> Upload -> Upload files

A screenshot of a computer

Description automatically generated

Load Tables

A screenshot of a computer

Description automatically generated

Once data is loaded to Tables you can use also use SQL Analytics Endpoint

1. Click on Open Notebook
2. New Notebook -> If you want to create a new Notebook.
3. Existing Notebook -> If you want to use Existing Notebook

You can use existing notebook option , and use DataEngineeringEconomicDataAnalysis Notebook

1. **Copilot for Data Engineering**

Click on Copilot , a prompt opens on left and a notebook cell is added to install required packages for Copilot

A screenshot of a computer

Description automatically generated

Once the Packages are installed it also shows commands that can be used to give instructions to copilot using Notebook

**Main Commands we used**

* %%code - Generate code to work with or visualize your data.
* %%fix\_errors - Fix errors in a cell.

You can read more about other commands in notebook

Step 1 :-

Type “Load companies into a DataFrame

A close-up of a message

Description automatically generated

Step 2 :- Use commands in cell to load economic\_data and purchase table to dataframe

%%code

Load economic\_data into DataFrame

%%code

Load purchase into DataFrame

Step3 :- Data Analysis

1. Mortgage trend by quarter

**%%**code

display graphical trend of mortgage\_rate\_30y based on quaterandyear order by year,quarter

1. Trend of CPI By Quarter

**%%**code

display graphical trend of cpi based on quaterandyear order by year,quarter

1. Trend of GDP Value by Quarter

**%%**code

display graphical trend of gdp\_value based on quaterandyear order by year,quarter

1. We also Tried to find relation between gdp\_value and unemployment rate

**%%**code

relation between gdp\_value **and** unemployment\_rate

1. Graphical trnd between GDP and unemployment rate

**%%**code

show graphical trend between gdp\_value **and** unemployment\_rate **with** gdp\_value **as** X **and** unemployment\_rate **as** Y **and** order by gdp\_value

1. Trend of CPI based on GDP

**%%**code

show graphical trend between gdp\_value **and** cpi **with** gdp\_value **as** X **and** cpi **as** Y **and** order by gdp\_value

**Machine Learning**

Create a **Linear Regression Model** Using Copilot to Predict Mortgage rates for next 3 quarters

**%%**code

Create a model to predict mortgage\_rate\_30y **for** future quaterandyear based on current data

Displayed the prediction in graph along with current mortgage rates

**%%**code

show above prediction **with** current data **in** a graphical way

We further combined economic\_date and purchases table to identify reveue for each productcategory for each quarter for each company

**%%**code

create a new table by joining economic\_data **and** purchase table

join on columns quarterandyear,quarter,year

sum revenue based on productcategory

table should have columns quarterandyear,quarter,year,company\_id,productcategory,revenue

We had errors but we did following changes

1. \*\*updated above code to include dataframes created above to resolve error
2. used %%fix\_errors remove duplicate columns in join to fix further error\*\*

Combined the table created above with company table to add company name

**%%**code

add a new column **in** df\_new\_table **as** company\_name

**with** company\_name **from** companies table based on company\_id

Saved this table as company\_sales table

**%%**code

save df\_new\_table\_with\_company\_name **as** table company\_sales

Calculated Total Revenue for Each Company

**%%**code

calculate total revenue **from** revenue **for** each company **from** company\_sales

**and** order by total revenue

rename sum(revenue) to Total\_Revenue

Identify company having top revenue

**%%**code

based on above which **is** the company having top revenue

Identify Top selling product of company

**%%**code

what **is** the top selling productcategory of 'Davis Group'

1. **Co-Pilot for Power BI**

1. Go to LakeHouse , Click Sematic Model , name the model **PowerBIReportingDataset**
2. Add tables economic\_data , company\_sales,companies
3. Open the Co-pilot prompt

It will give you report suggestions

We created first report from its suggestion -> Revenue and Product Category Comparison

We created second report from prompt -> Company Performance Analysis

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. **Co-pilot for Data Factory**

3rd feature for Data Factory we looked to explore

1. Open Data Factory
2. Click on New DataFlowGen2
3. We connected to Lakehouse and selected Tables Company\_sales
4. We copied the table as company\_sales\_upload
5. **We deleted a column for Quarter using Copilot for Datafactory prompt**
6. We added a Data Destination as Azure SQL Database
7. We published the dataflow as DataFlowCoPilotLakeHouseToSQLDB

A screenshot of a computer

Description automatically generated

We verified the data exists in SQL Server instance.