

دانشگاه حکیم سبزواری

دانشکده مهندسی برق و کامپیوتر

مستند سازی پروژه

Microsoft
Power BI/Fabric

درس: شبکه های کامپیوتری

دانشجویان: سیده بیتا امیری – عارفه عمیدیان

استاد مربوطه: خانم دکتر مینا ملک زاده

فروردین ماه 1403



فهرست

4.....	مقدمه
5.....	Microsoft Power BI چیست؟
6.....	کاربردهای Microsoft Power BI
7.....	نصب Microsoft Power BI
17.....	آشنا شدن با محیط Power BI
21.....	First dashboard in Power BI
22.....	تنظیم اولیه در Microsoft Power BI
27.....	پیاده سازی پروژه در Microsoft Power BI
27.....	مرحله اول: Getting Data
32.....	مرحله دوم: Clean & Transform Data
39.....	مرحله سوم: Design a Model
42.....	مرحله چهارم: Making Visuals
49.....	Microsoft Fabric چیست؟
66.....	ورود به وبسایت Microsoft Fabric
77.....	آشنایی با محیط Microsoft Fabric
89.....	Lakehouse
91.....	پیاده سازی پروژه در Microsoft Fabric
106.....	مقایسه Microsoft Fabric & Microsoft Power BI
108.....	منابع و تشکر

مقدمه

همه‌ی سازمان‌ها دارای حجم عظیمی از داده‌ها هستند. اما داستان‌های پشت این داده‌ها چیست و بر اساس آن چه تصمیماتی باید بگیرید و چگونه به تحلیل و بررسی آنها بپردازید.

با دو ابزار به کمک ما آمده:

Microsoft Power BI .1

Microsoft Fabric .2

در ادامه به بررسی و پیاده سازی یک دیتابست با نام Social networks می‌پردازیم.

*لازم به ذکر است Power BI در لپتاپ خانم امیری و Fabric در لپتاپ خانم عمیدیان انجام می‌شود.

Power BI



چیست؟ Microsoft Power BI

اینجا میخوایم یک معرفی داشته باشیم که کدوم نسخه از Power Bi رو نصب بکنیم و کارمون رو شروع بکنیم.

Converting Data into Insights



به طور کلی Power BI او مده تا از منابع داده‌ی مختلف، داده‌ها را جمع کنه و بعد اونها رو به صورت گزارش در اختیار شما بذاره تا بتونید بر اساس اون بینش بدست بیارید و تصمیم گیری کنید.

کاربردهای Microsoft Power BI

حالا ابزارهایی که برای این مسیر طراحی کرده اینا هستند:

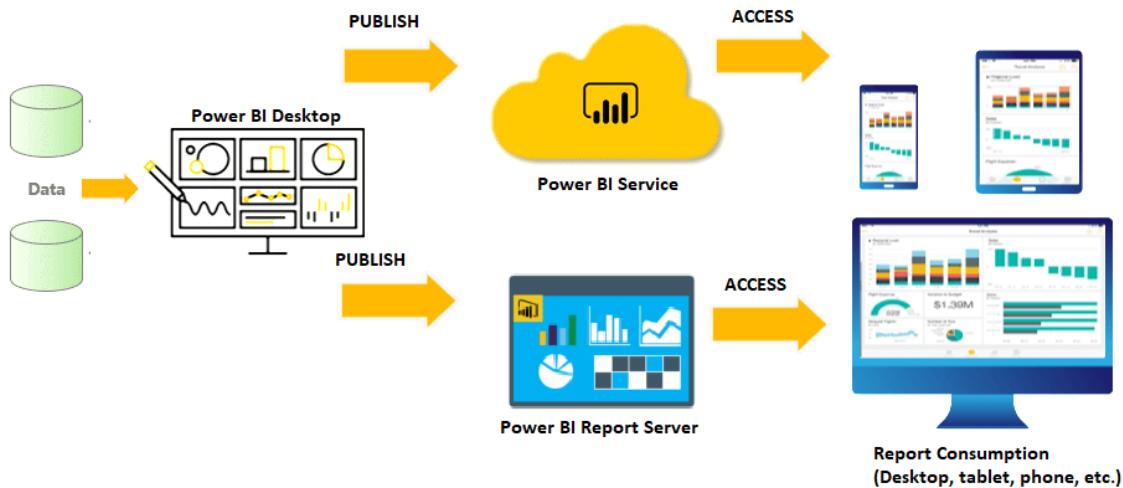
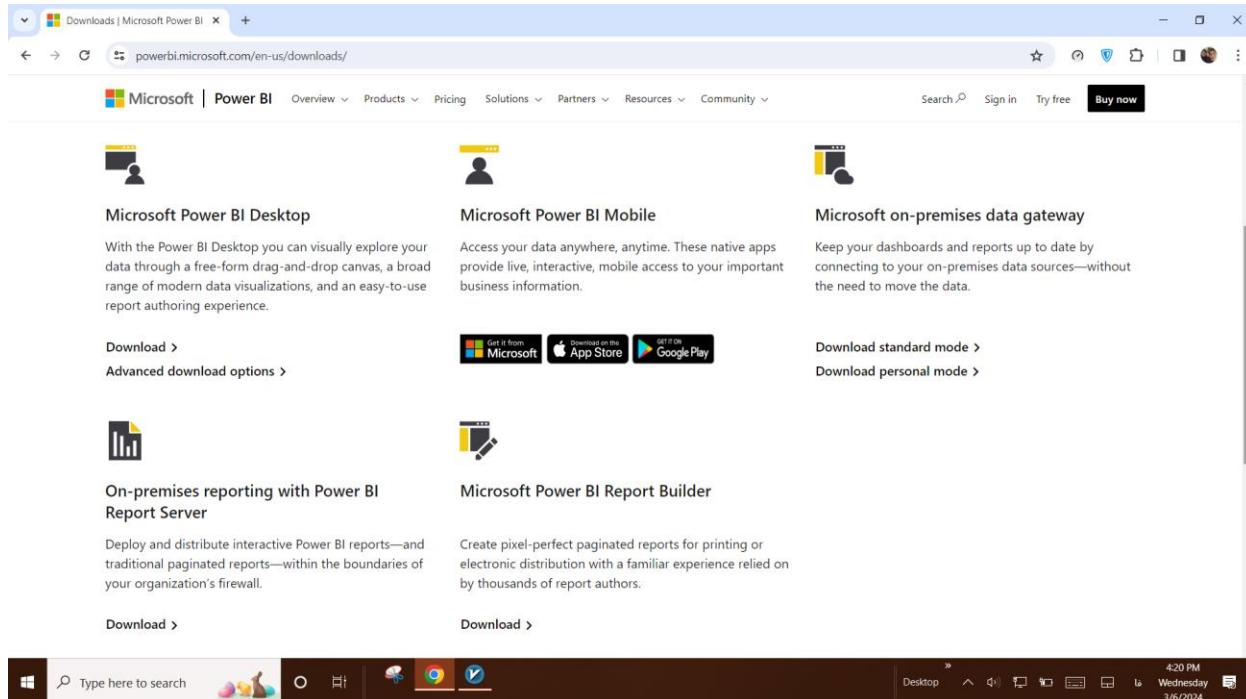
1. اول از همه او مده نرم افزار **Power BI Desktop** را برای شما ساخته که این نرم افزار محیط طراحی و محیط اصلی کار ما هست. این نرم افزار فقط روی سیستم عامل ویندوز نصب میشه و روی لینوکس و مکناتش نمیتونیم نصبش کنیم. بعد از اینکه ما گزارشات رو آماده کردیم اصطلاحا باید اونها رو publish یا منتشر کنیم.

2. وقتی گزارشات publish میشن، میرن روی یک server ی قرار میگیرن که دیگران هم بتونن به اون دسترسی داشته باشند. ما میتوانیم روی **Power BI Service/ Power BI Report server** اونها رو قرار بدیم. همه ی این سرویس ها به صورت تحت وب در دسترس هستند و فرقی نمیکنه شما از چه سیستم عاملی اونها رو وارد بکنید.

3. مکناتش، لینوکس، موبایل یا هر چیز دیگه ای باشه اما یک ابزار اختصاصی دیگه هم مایکروسافت داده به نام **Power BI Mobile**. اپلیکیشن موبایل Power BI به راحتی روی گوشی های IOS, Android نصب میشه و شما میتوانید نمای موبایل گزارشاتون رو به زیبایی از طریق اپلیکیشن های موبایل به اشتراک بگذارید.

نصب Microsoft Power BI

اگر یک سرج ساده توی اینترنت کرده باشید که چه نسخه ای از Power BI رو نصب بکنیم با نسخه های متفاوتی روبرو میشید.

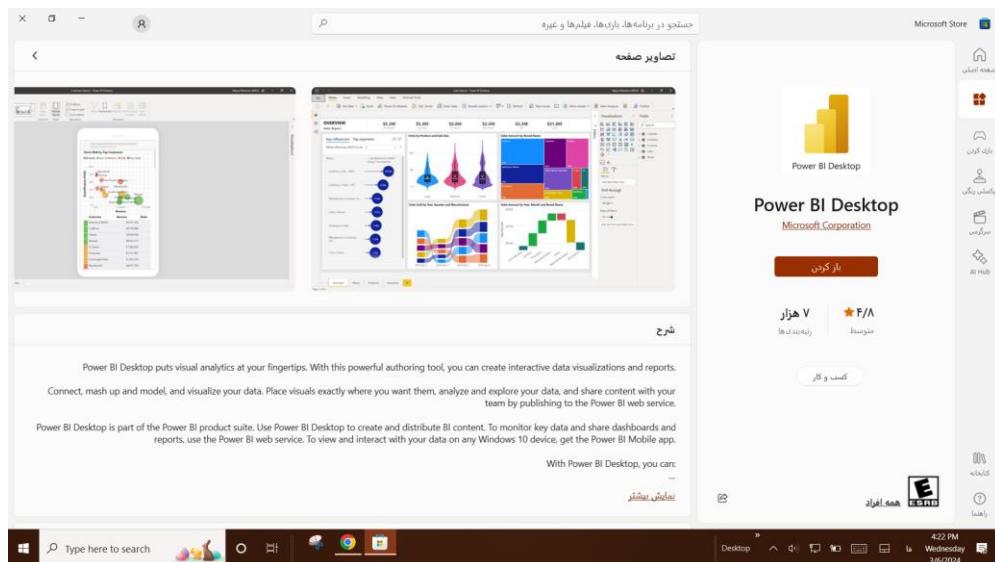


خب ما قراره نرم افزار Power BI Desktop رو نصب بکنیم. بنابراین بین نسخه های مختلف Power BI Desktop باید بدلونیم که دو نسخه ای اصلی برای Power BI Desktop وجود داره:

- . یکی Power BI Desktop عادی که برای ارسال داده و انتشار روی Power BI Cloud طراحی شده.
- . دویکی Power BI Report Server که برای انتشار روی سرورهای Report Server طراحی شده.

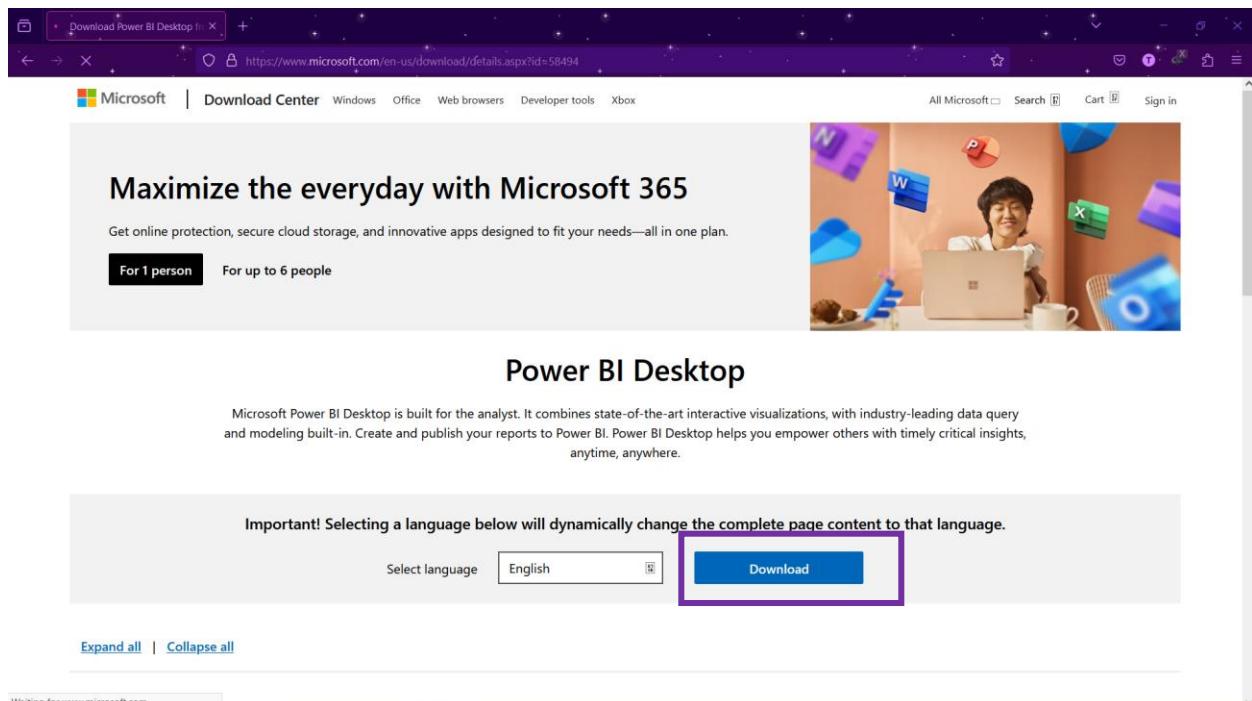
این دو تا با هم یک تفاوت هایی دارن اما میایم و نسخه‌ی Power BI Desktop را برای آموزش انتخاب میکنیم و این نسخه را نصب میکنیم.

نرم افزار Power BI رایگانه و روی سایت خود Microsoft بدون هیچ محدودیتی در دسترسه و میتوانید اون رو دانلود کنید. از طریق Microsoft Store هم امکان دانلود هست.

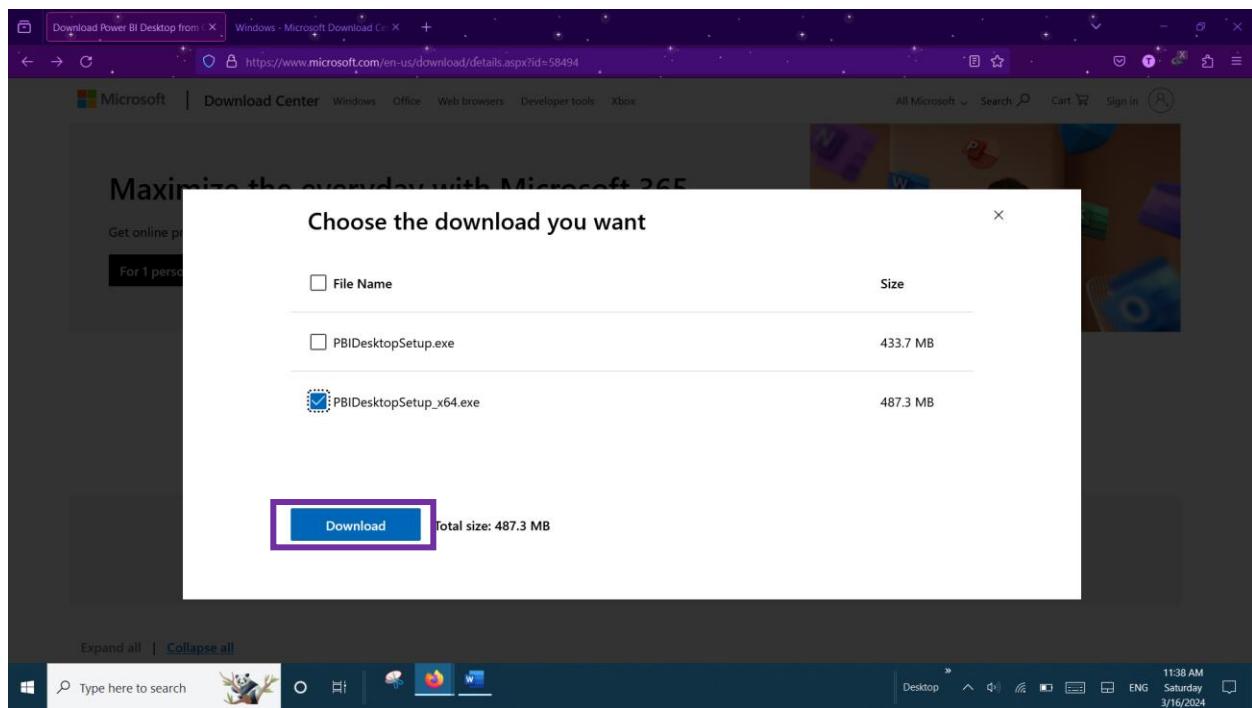


برای دانلود از سایت ابتدا گوگل میکنیم. بین نتایجی که پیدا میشه اونی که سایت Microsoft مستقیم هست رو انتخاب بکنید.

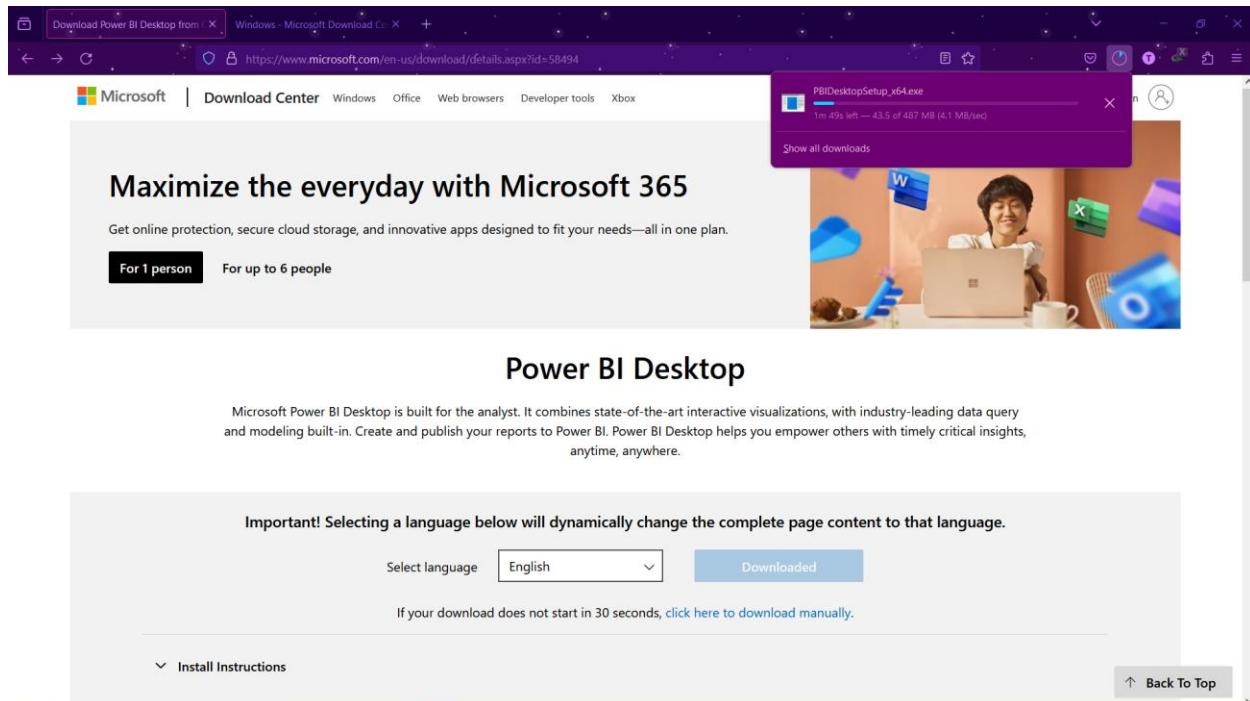
از سایت Microsoft میتوانید نسخه ی 32 بیتی یا 64 بیتی رو دانلود بکنید. و توی صفحه ی دانلود روی کلید Download کلیک بکنید.



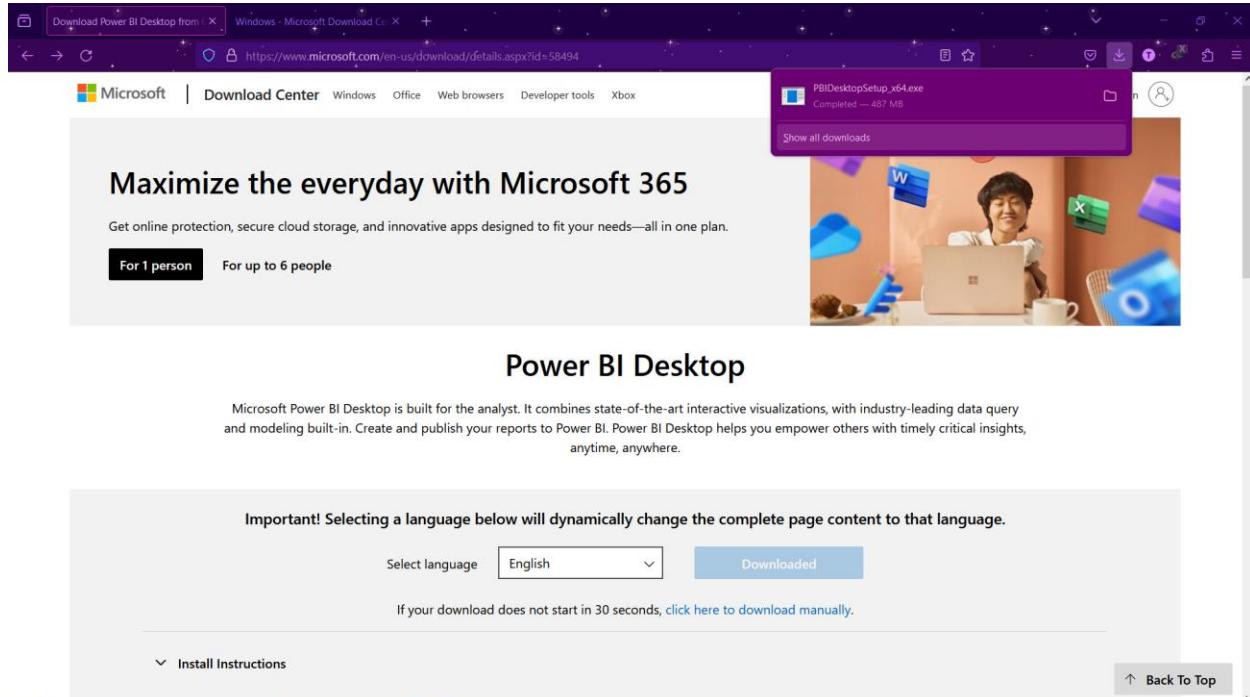
اینجا میتوانید انتخاب کنید که بسته به سیستم عامل خودتون نسخه ی 32 بیتی یا 64 بیتی دانلود بشه.



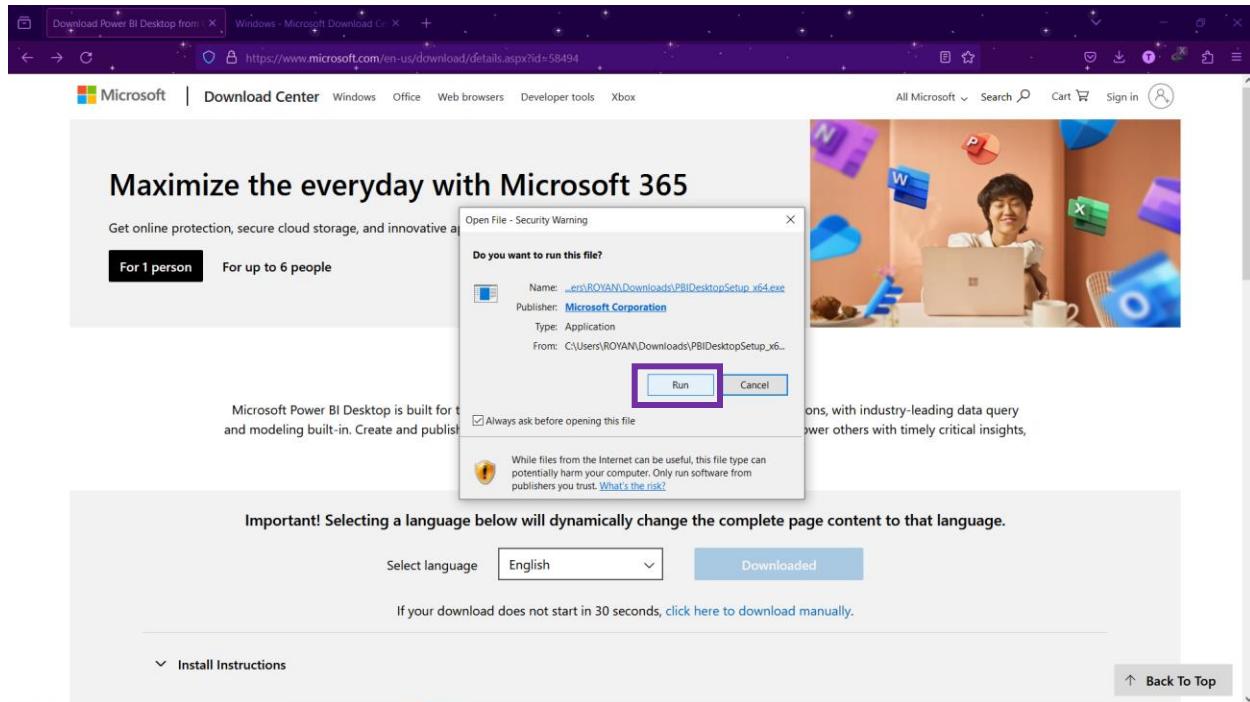
نرم افزار کمتر از 500MB هست و نرم افزار سنگینی هم به حساب نمی‌آید. با یک اینترنت معمولی کمتر از ده دقیقه طول می‌کشه.



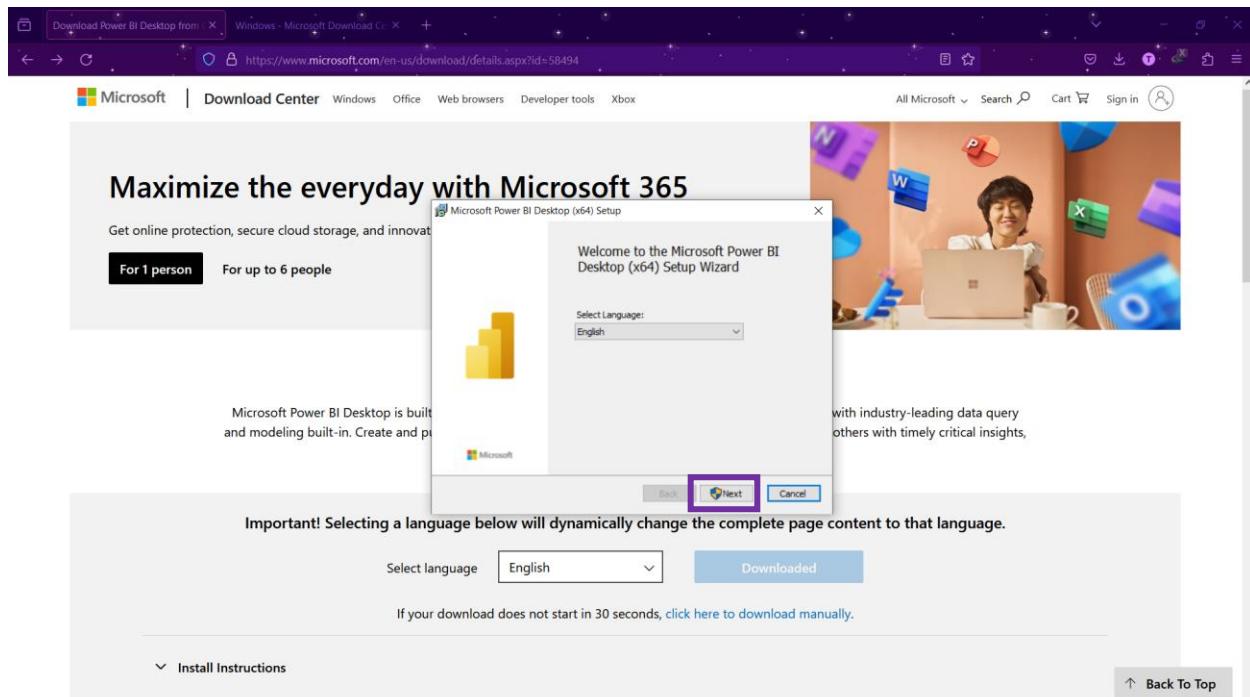
خب دانلود کامل شد.

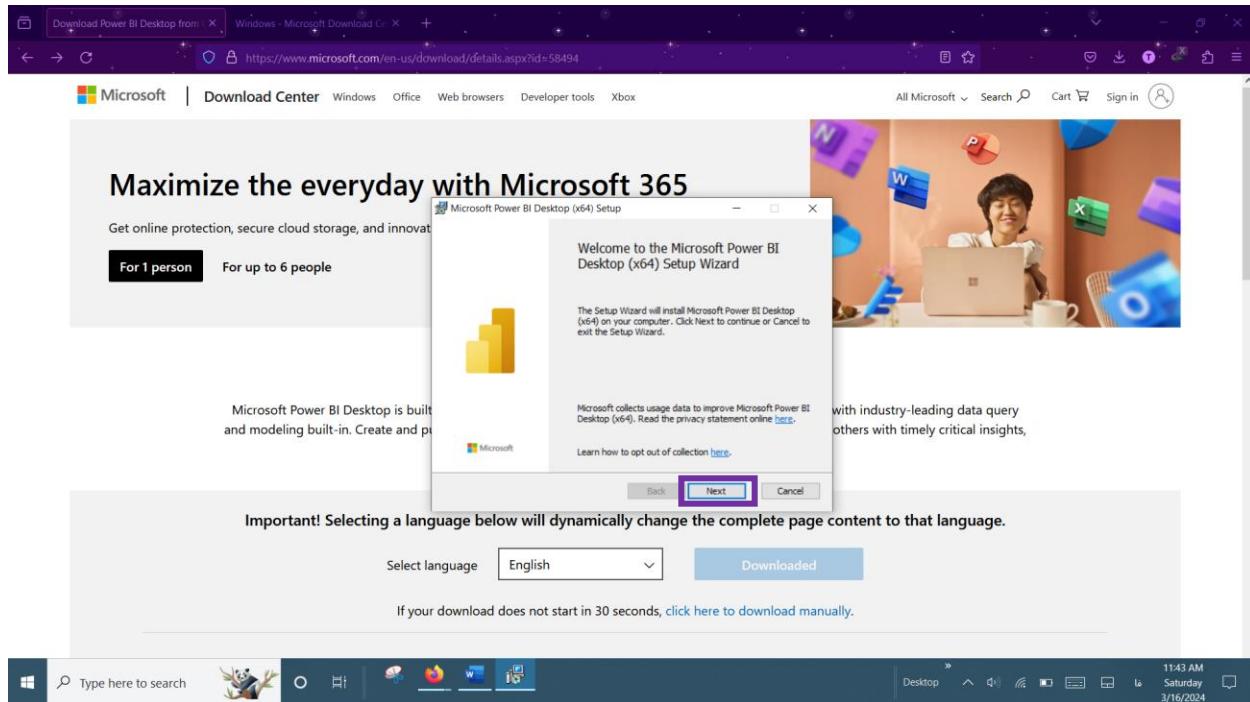


اجرای میکنیم.

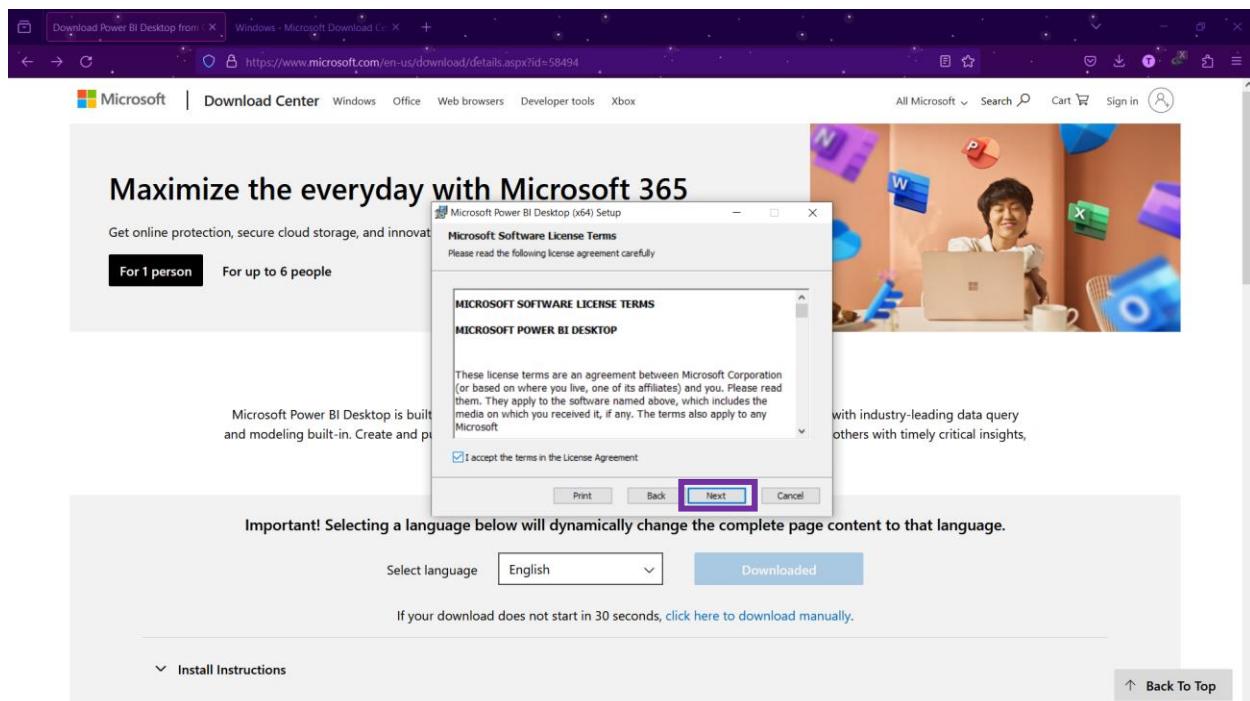


زبان default رو انگلیسی بذارید باشه. Next رو میزنیم.

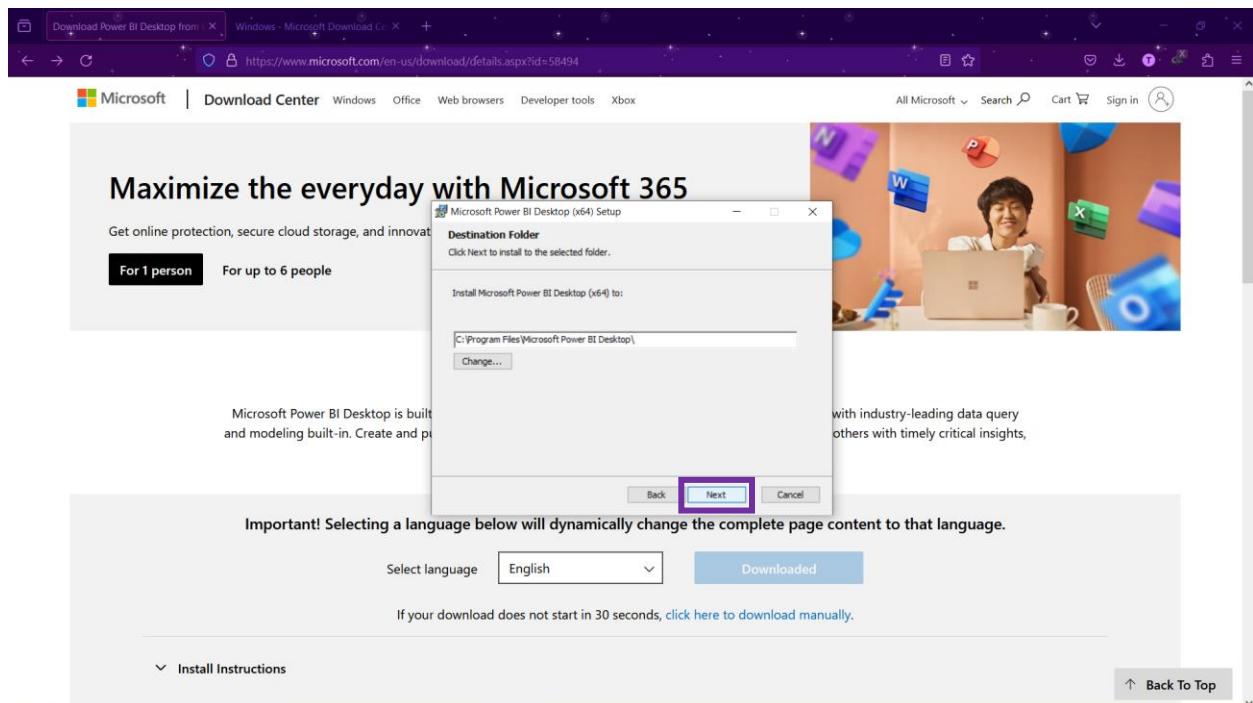




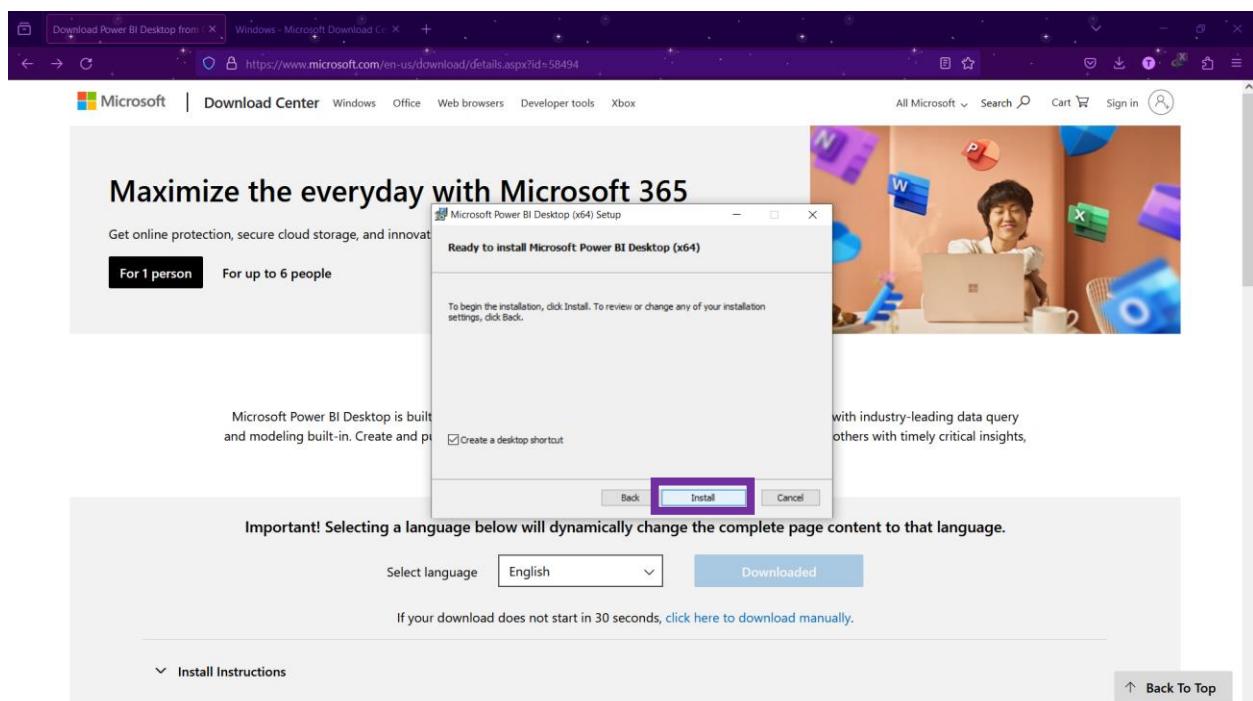
.Next Microsoft خواسته های کنید رو. بعد Accept



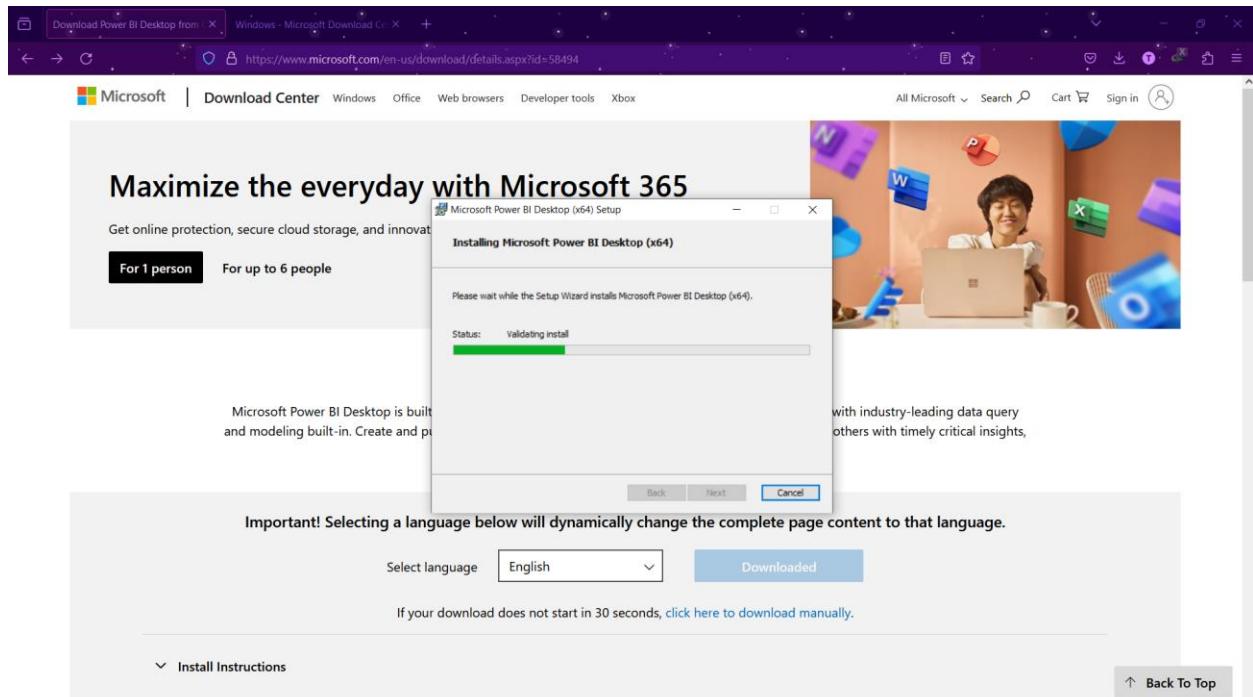
پیشنهاد میکنیم روی درایوی نصب کنید که SSD داشته باشد اونجوری یک مقدار سریعتر کار میکند.



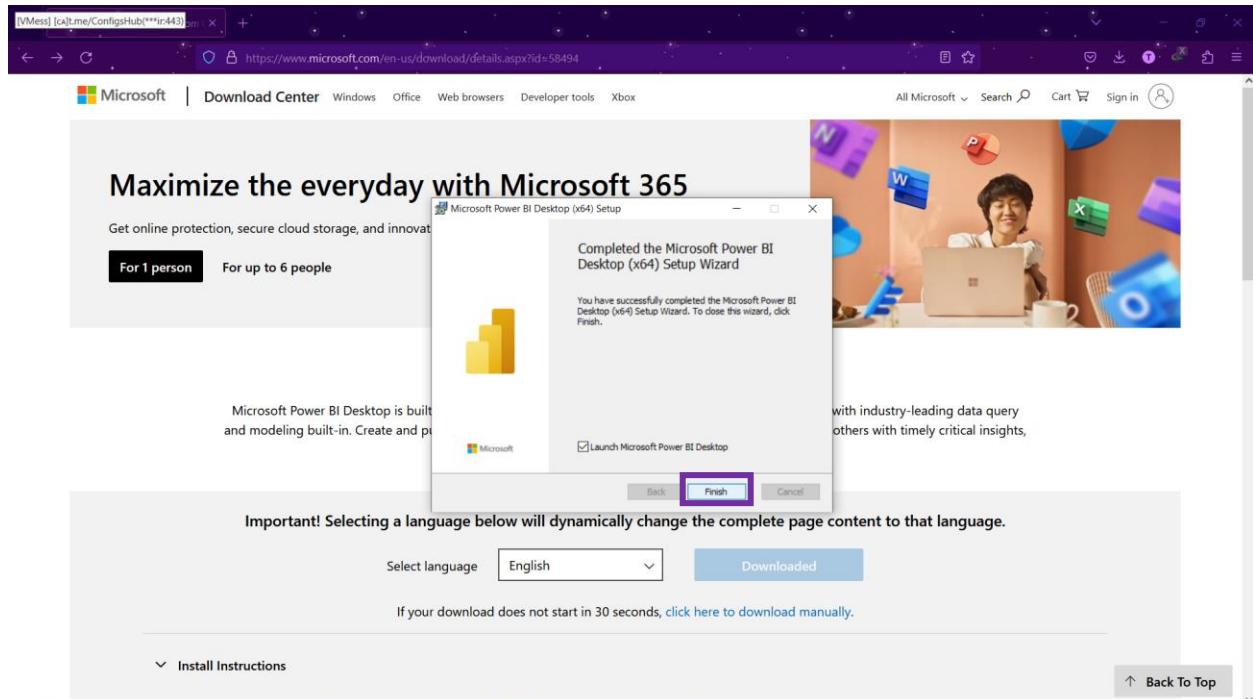
همرو انتخاب میکنیم و کلید install رو میزنیم.



زمانیکه تحلیل انجام میده رم بیشتری احتیاج داره. بنابراین اگر سیستمتون رم خوبی داشته باشه خیلی سریع کار میکنه.

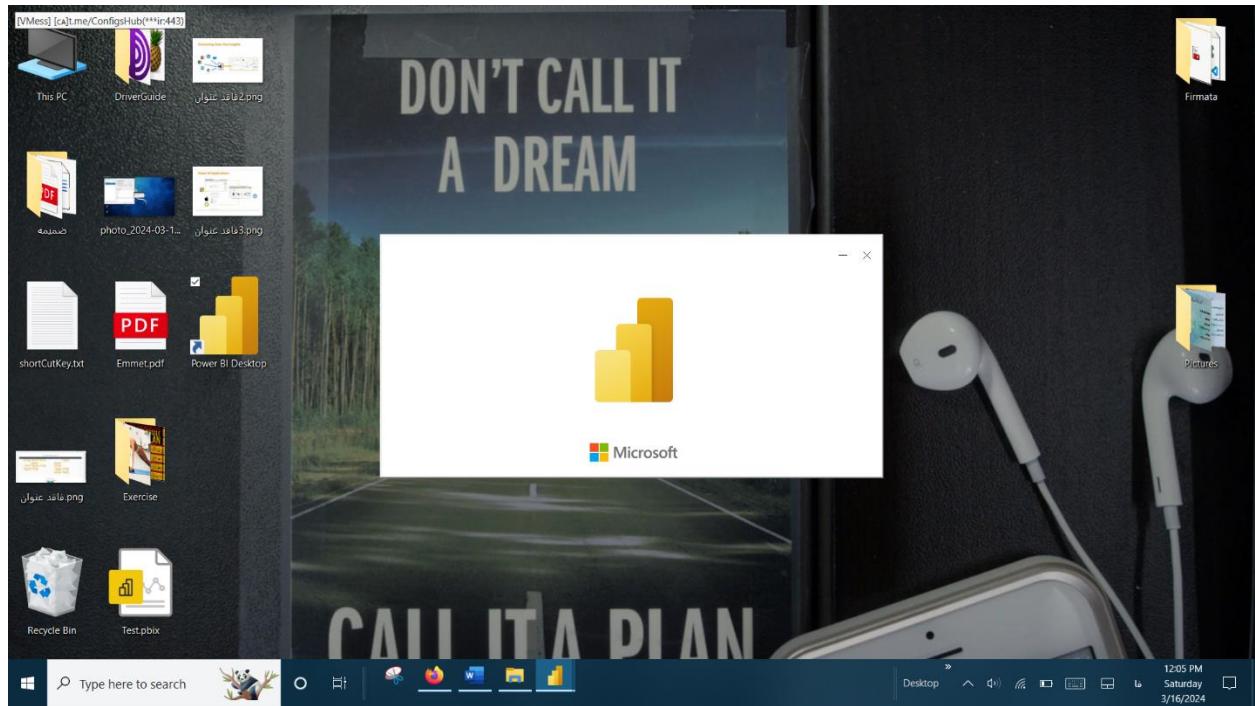


اینجا هم کافیه که Finish رو بزنید.

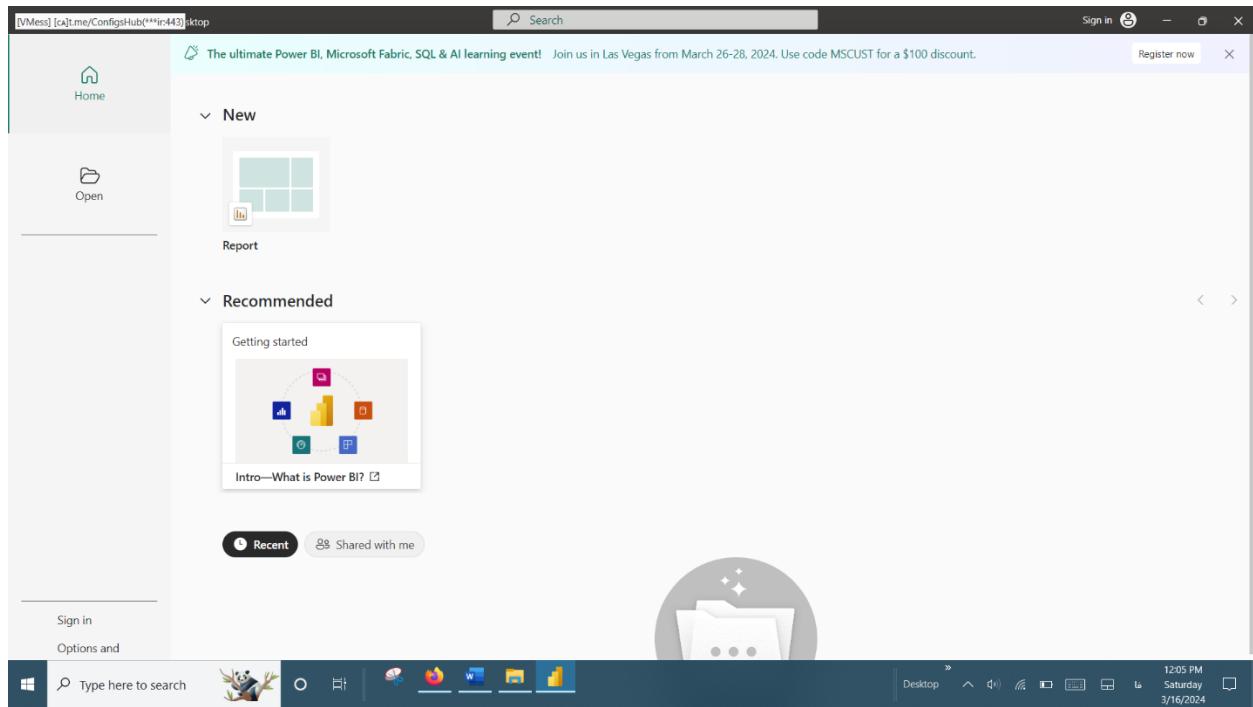


ما نصب شد.

بریم سراغ اجرای Power BI و ورود به دنیای جذاب Power BI Desktop



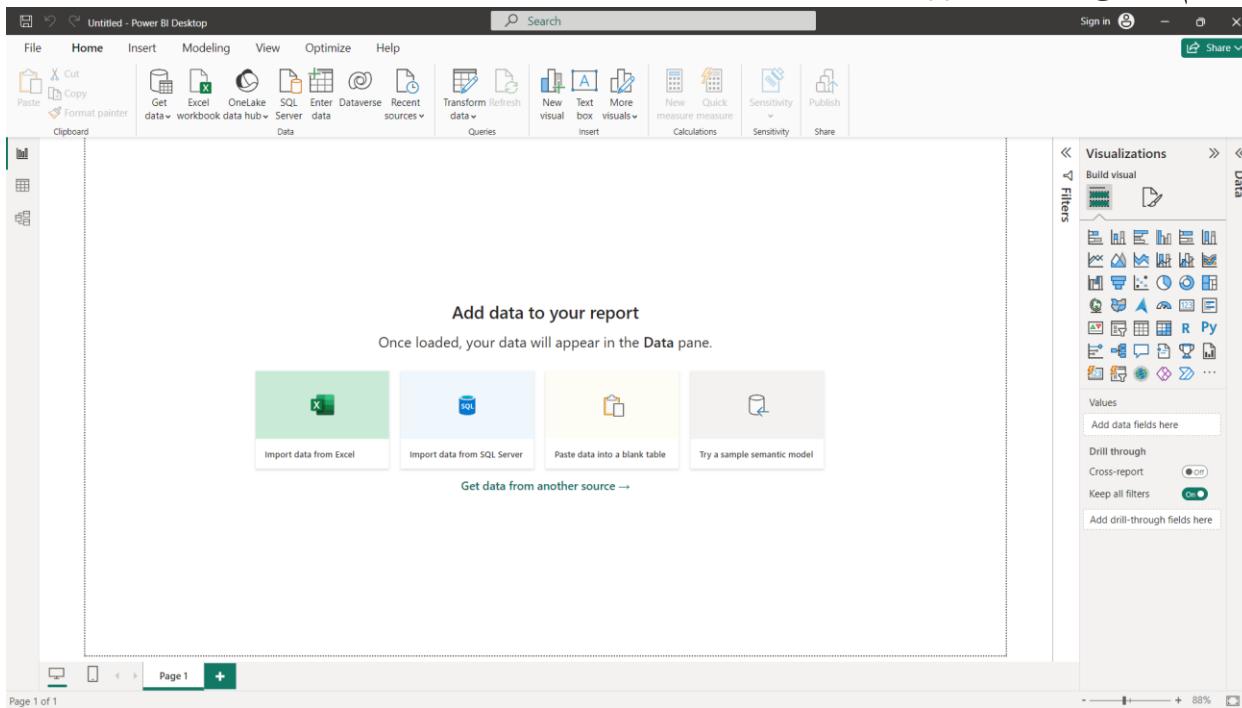
نرم افزار Power BI بدون هیچ دردرسی روی سیستم شما نصب میشے و محیط کار اون آماده ی شروع به کار شما میشه.



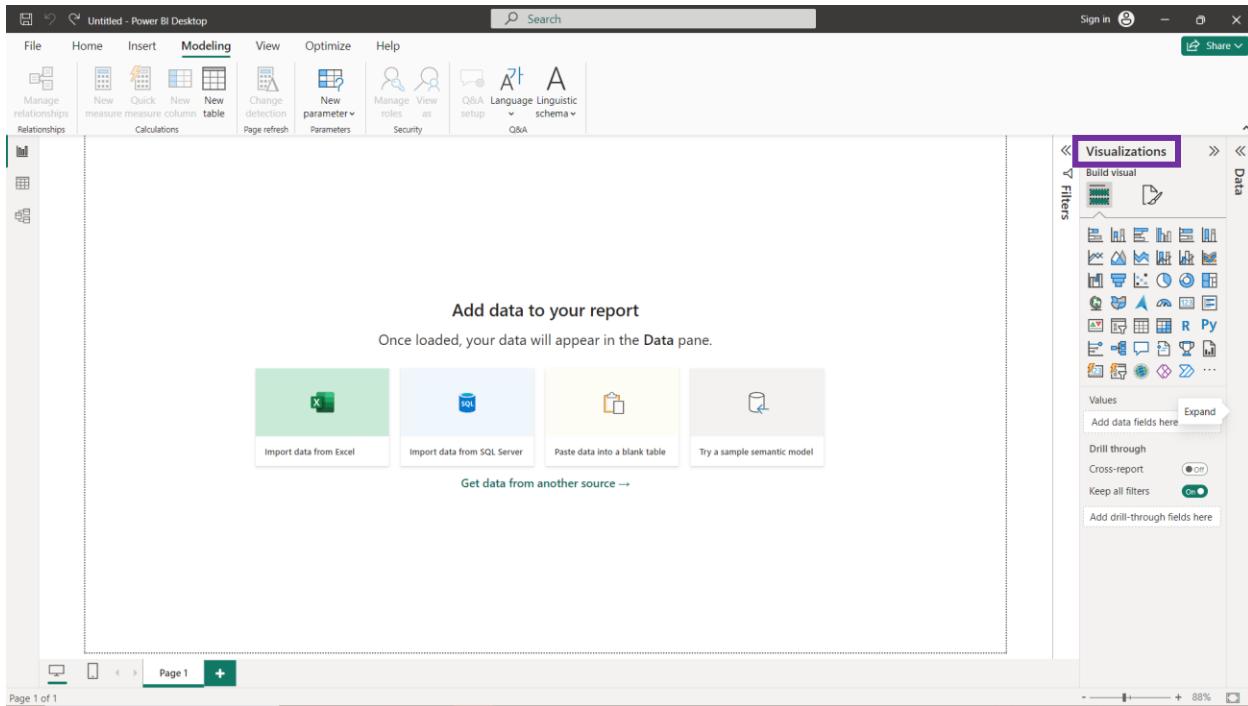
حالا میخوایم با محیط Power BI آشنا بشیم.

آشنایی با محیط Microsoft Power BI

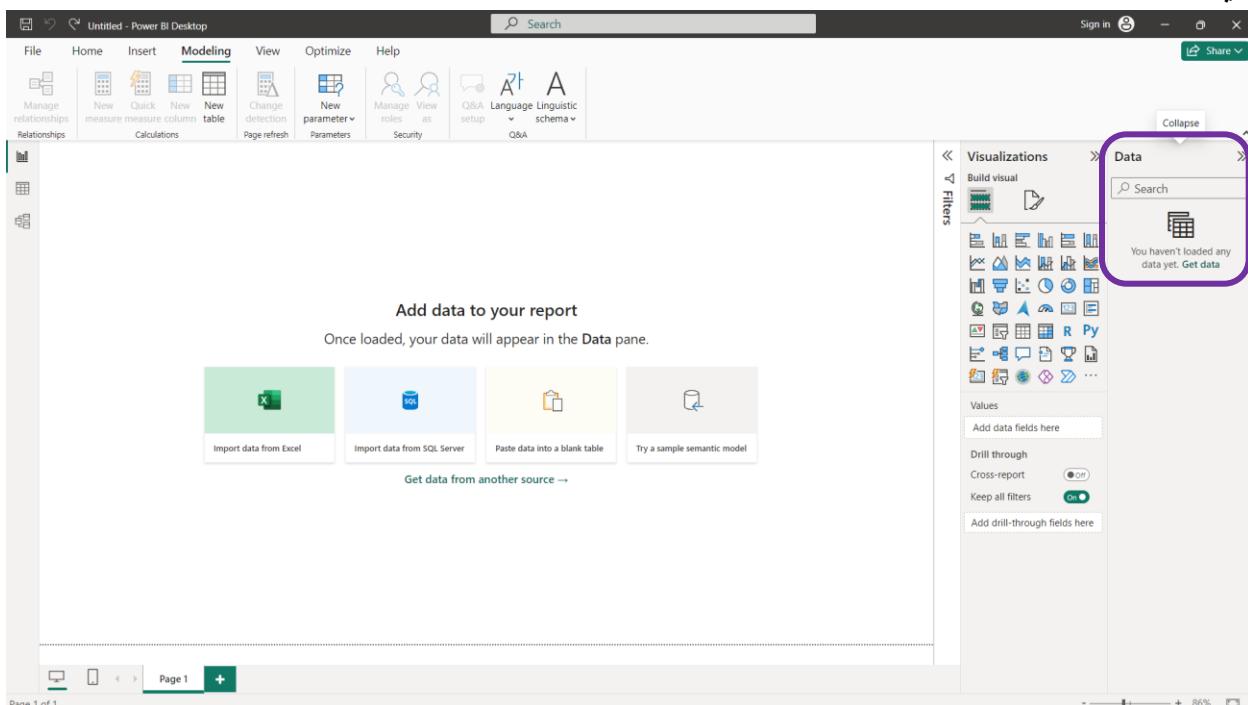
خب در این قسمت میخوایم بیانیم چه رو کجا میتوانیم پیدا کنیم. وقتی وارد Power BI میشیم و اون منوی اول رو میبینیم محیطی شبیه به این رو شما مشاهده میکنید.



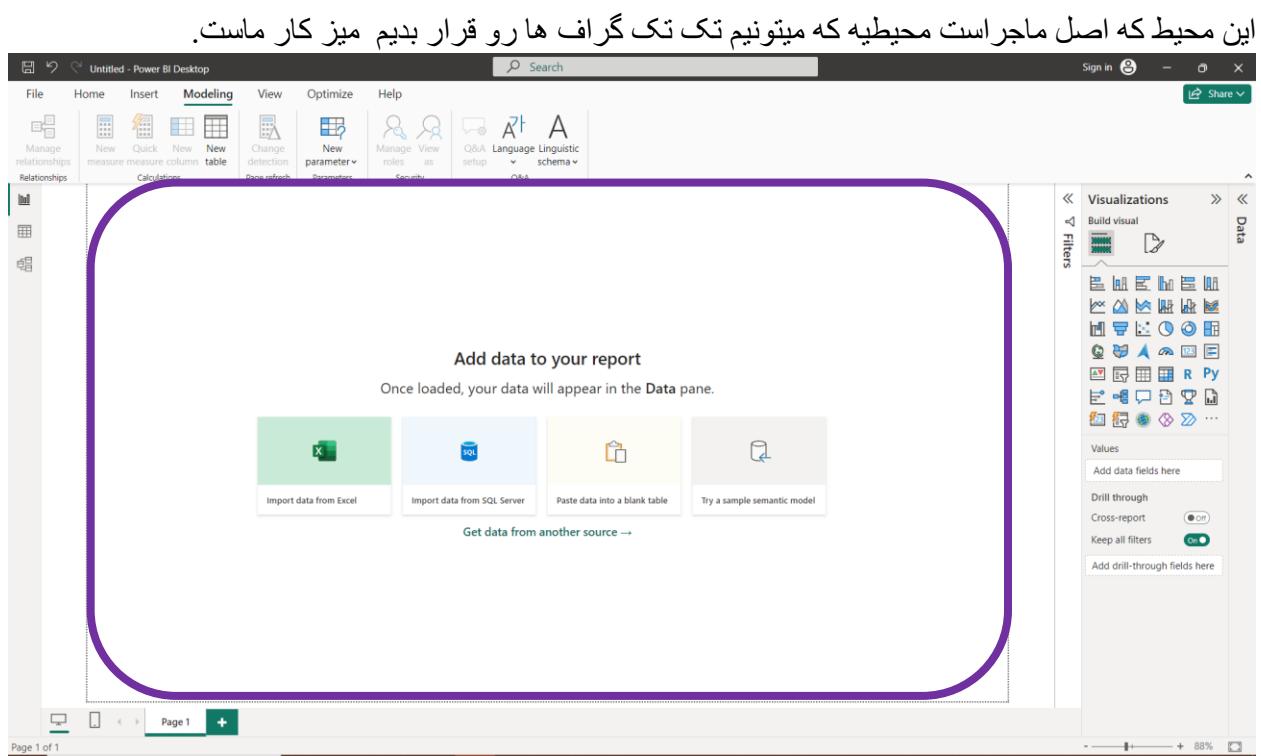
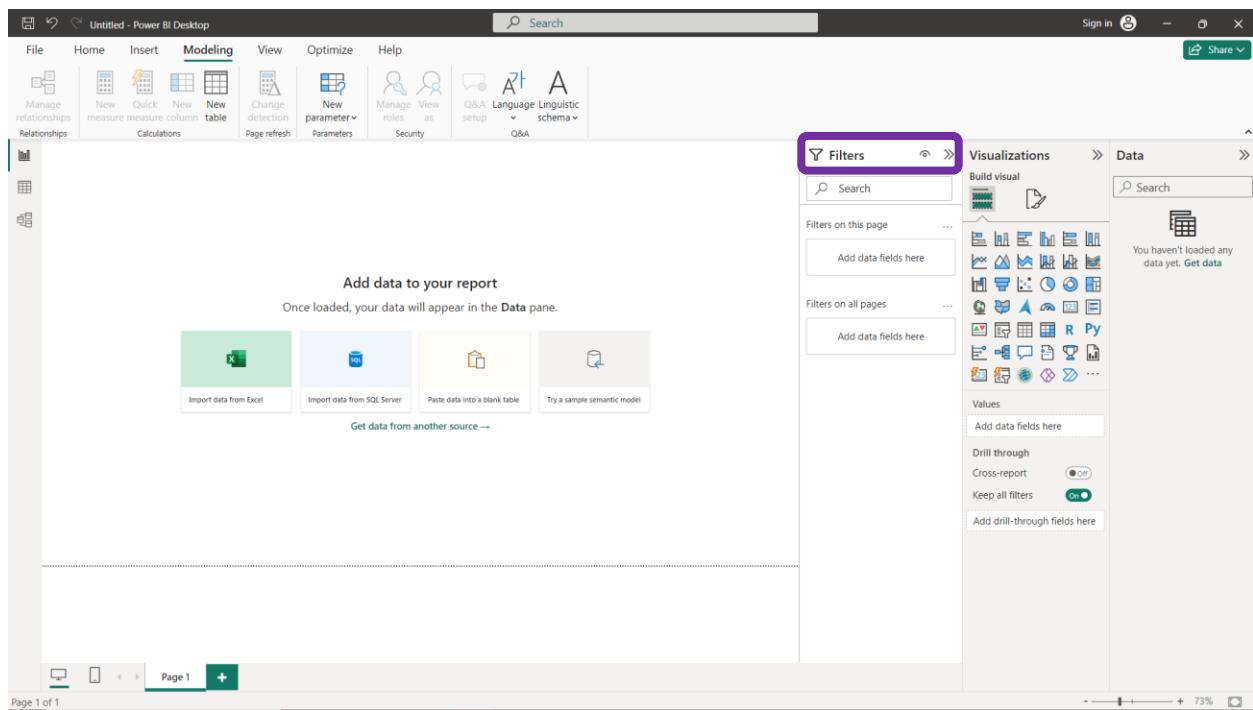
در ریبون یک قسمتی هست به نام table tools یک سری ابزار هست که میشه کارهای جورواجوری انجام داد. سمت راست یک بخش Visualizations داریم که همون گراف ها، نمودارها و ابزارهایی که شما نیاز دارید تا بتونید نمودارهای جذابتون رو بسازید اینجا قرار گرفته.



سمت راستش یک منوی کشویی قرار گرفته به نام Fileds/Data اینجا میتوانید اون داده هایی که از اکسل یا هر منبع داده‌ی دیگه‌ای رو اوردید رو پیدا کنید. از اینجا بردارید و یکیشو بندازید داخل صفحه و ازشون استفاده کنید.

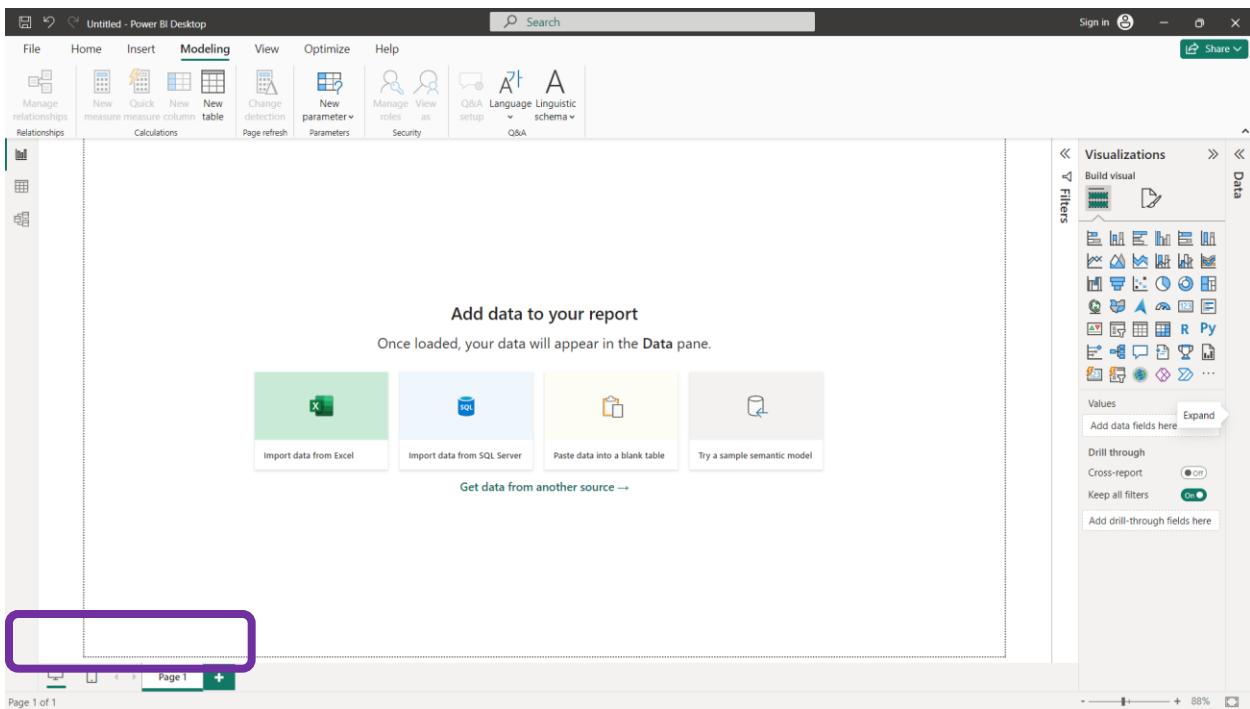


ی منوی دیگه داریم اینجا به نام Filters که شما میتوانید گزارشاتتون رو براساس فیلتر هایی که اینجا میدارید محدود بکنید.

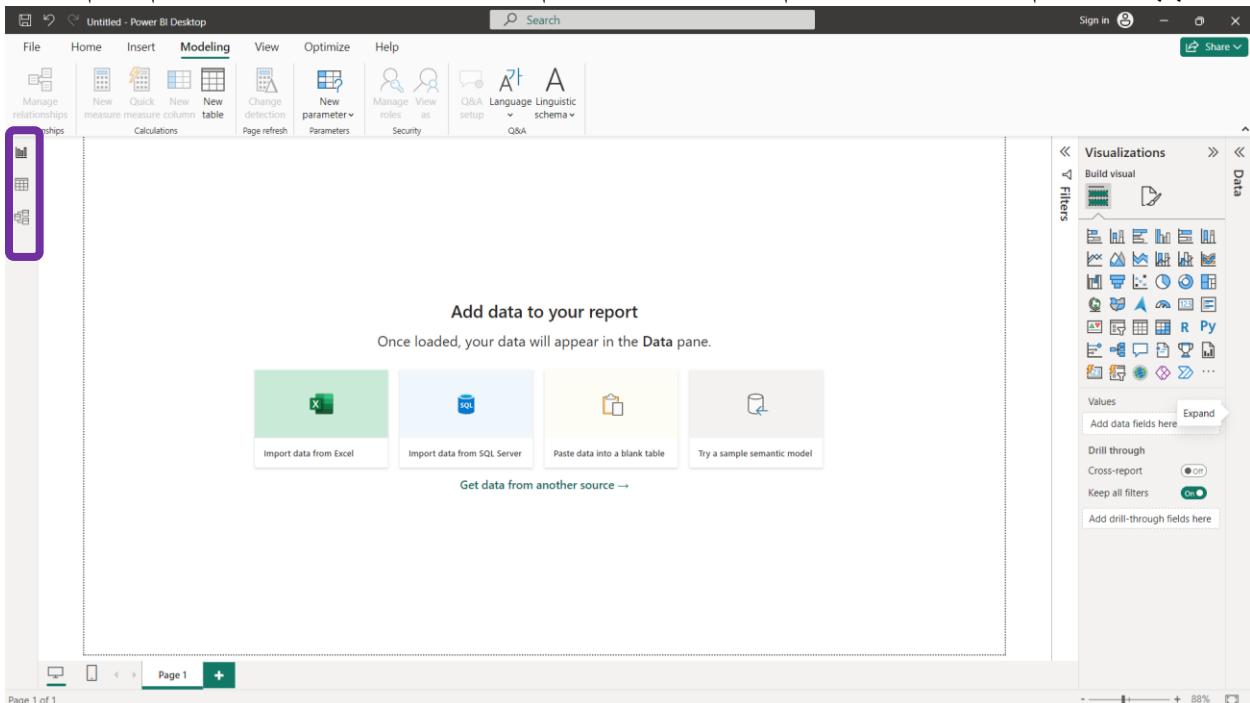


این پایین ما میتوانیم page های مختلفی از گزارش را بسازیم همون مدلی که توی اکسل هم شما احتمالاً اون رو

دید.

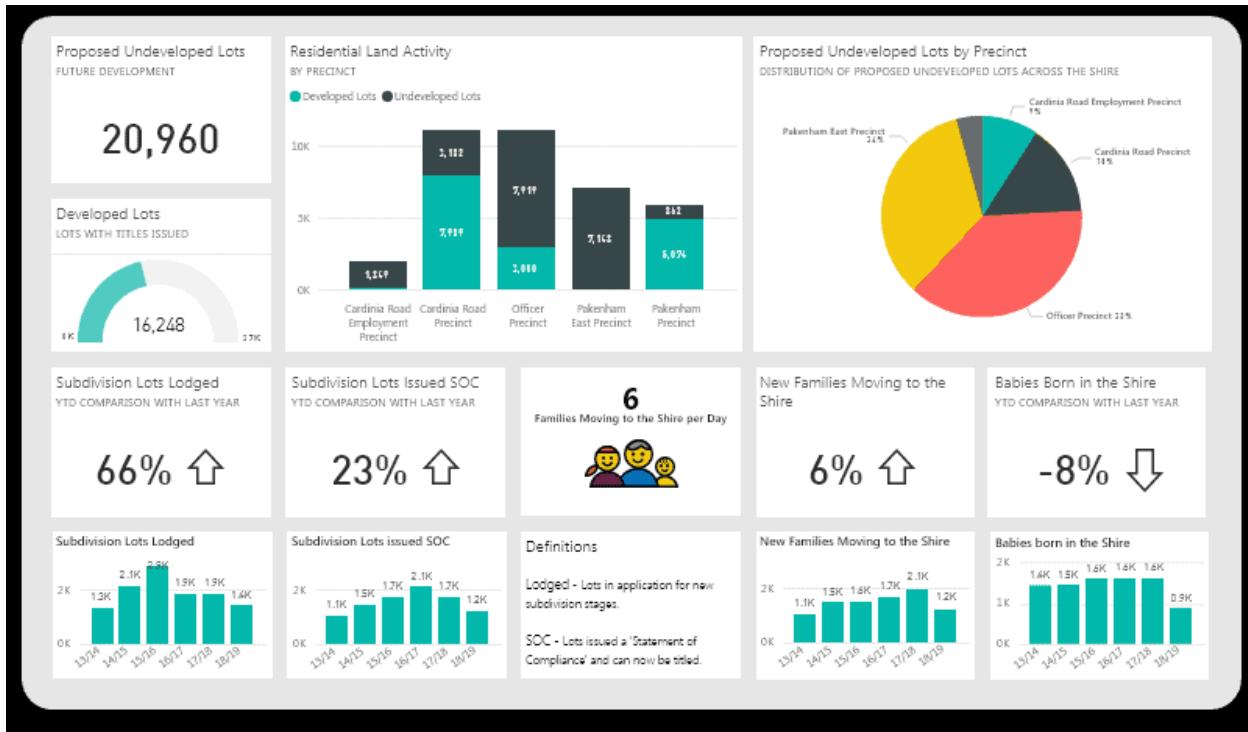


سمت چپ ما میتوانیم گزارشاتمون رو ببینیم، جداول داده رو ببینیم یا ارتباط اونها رو با همیگه ترسیم بکنیم.



خب بریم سراغ ساختن اولین داشبورمون.

First dashboard in Power BI



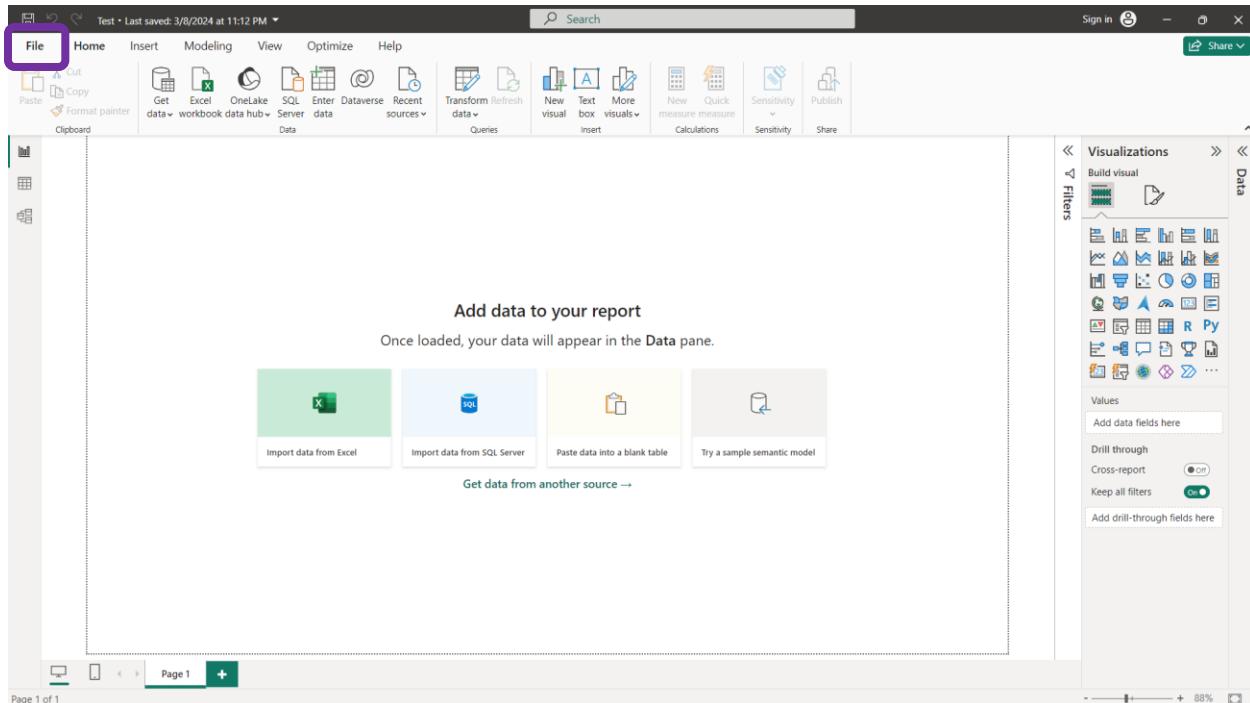
ما توی Business های مختلف با گزارشات سر و کار داریم. گزارشاتی که در یک لحظه تهیه میشند و ثابت هستند. اما Dashboard ها صفحاتی هستند که گزارشات پویا در اونها نمایش داده میشند. به این معنی که هم این گزارش ها به صورت زنده در دسترس قرار میگیرند و هم اینکه شما با کلیک روی بخش های مختلف گزارش ها، میتوانید فیلتر های زیبایی رو روی صفحه داشته باشید. پس از این به بعد با کلمه *Dashboard* ما بیشتر کار داریم و به مفهوم گزارشات زنده و پویایی هست که توی Power BI ایجادشون میکنیم.

قطعاً چیزی زیادی باید یاد بگیریم که بتونیم این Dashboard ها رو باهم دیگه با اطلاعات کامل بسازیم اما یک نکته ای که وجود داره سریع میریم سراغ کاربرد و ببینیم که نتیجه ای کارمون در آینده به چه شکل خواهد شد؟

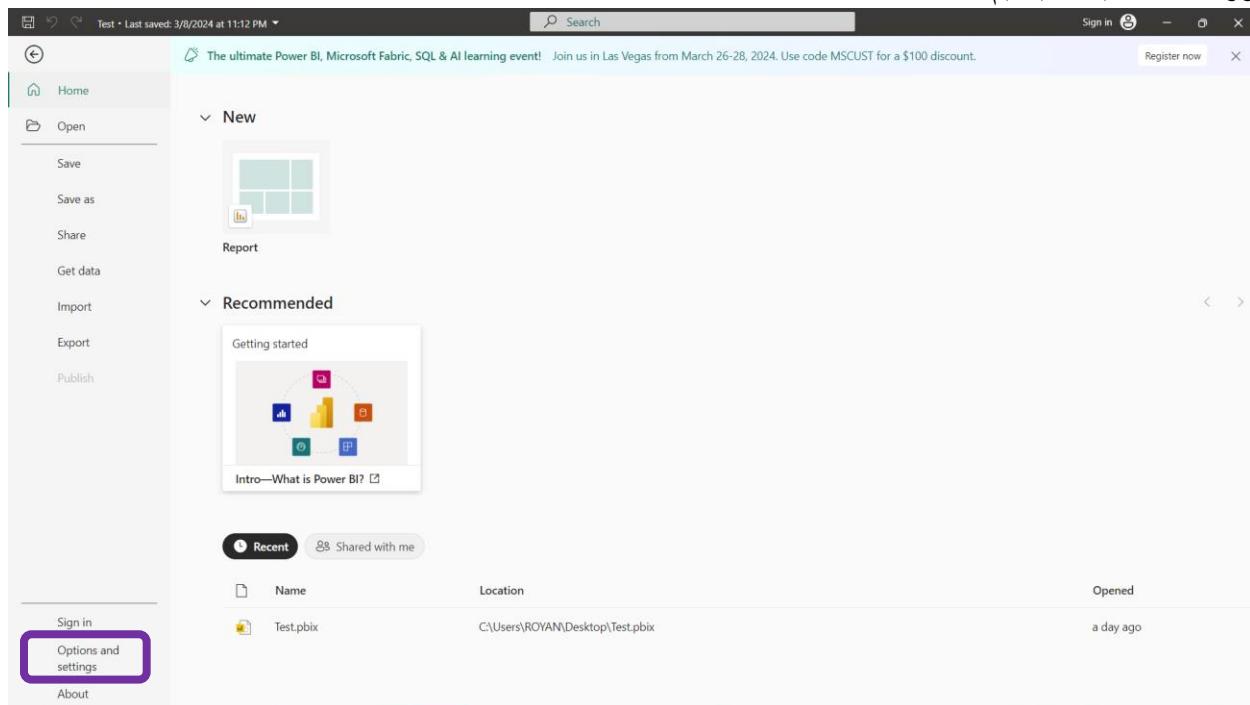
پس یک به یک به ترتیب کارها رو باید انجام بدیم به صورت مختصر توضیح میدیم اما بعداً قراره تک تک جزئیات گفته شده رو بررسی کنیم.

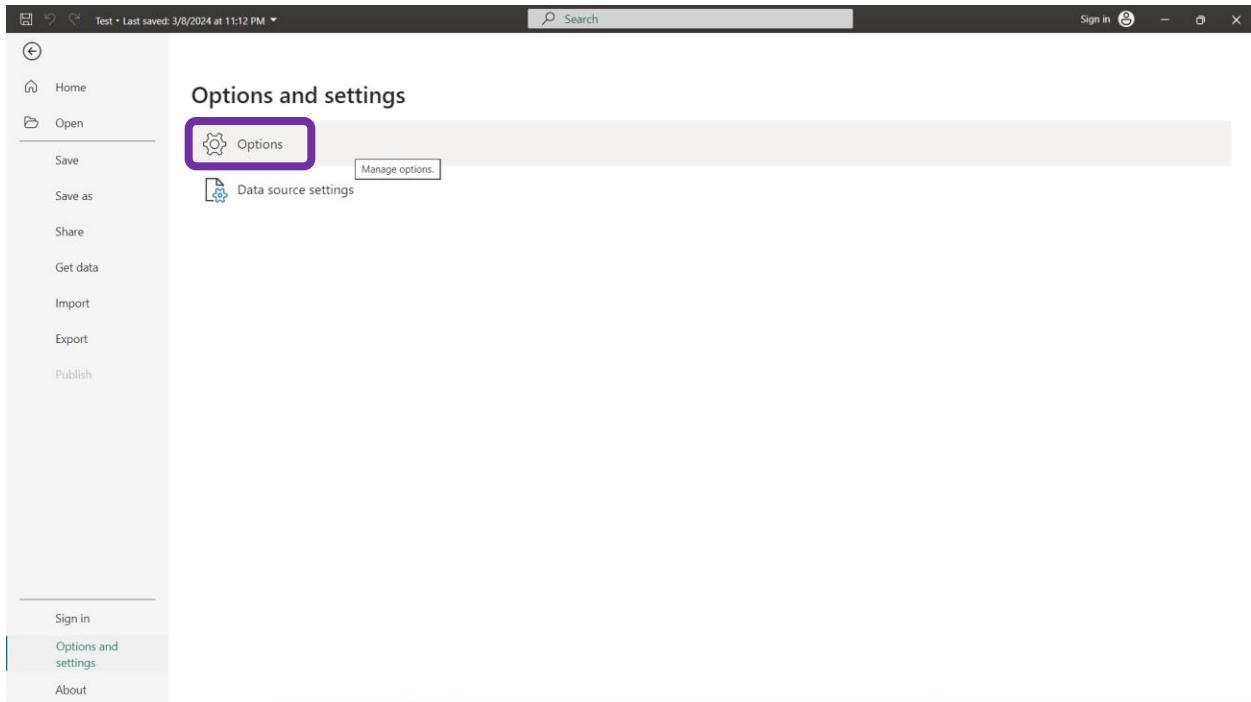
تنظیم اولیه در Microsoft Power BI

یک تنظیم اولیه برای استفاده از ویژوال‌ها باید داخل Power BI باید انجام بدم.

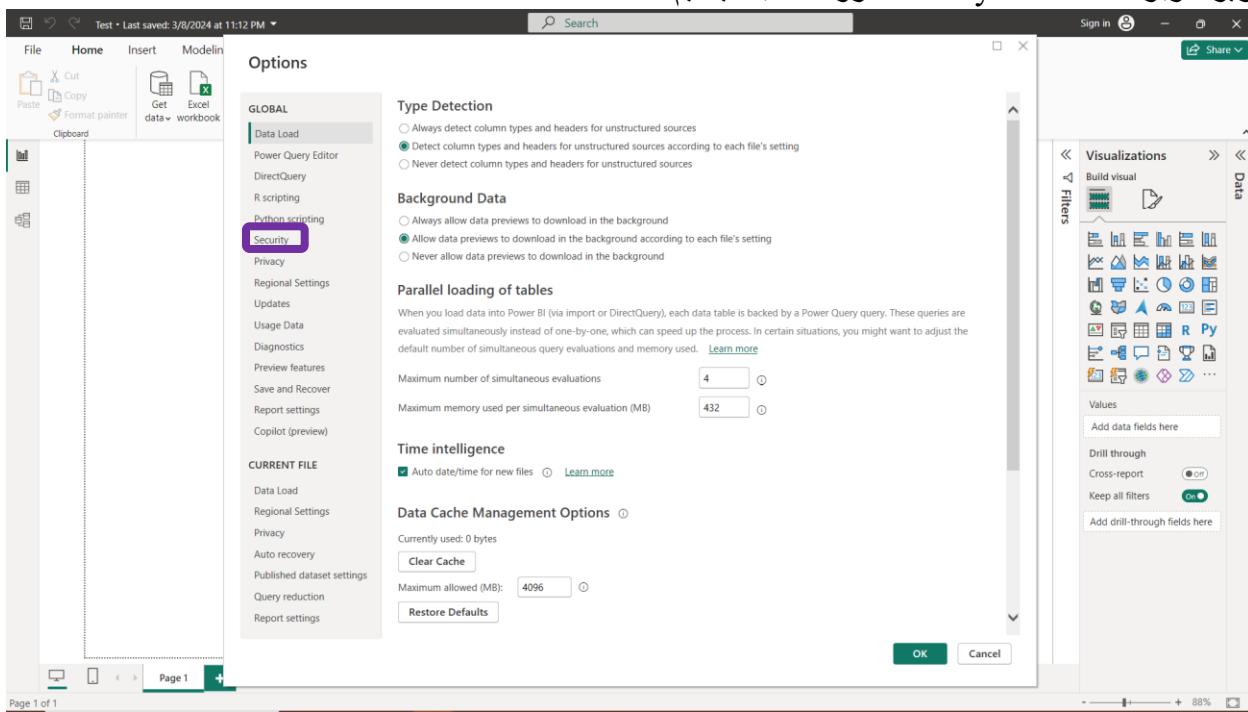


روی File کلیک می‌کنیم.

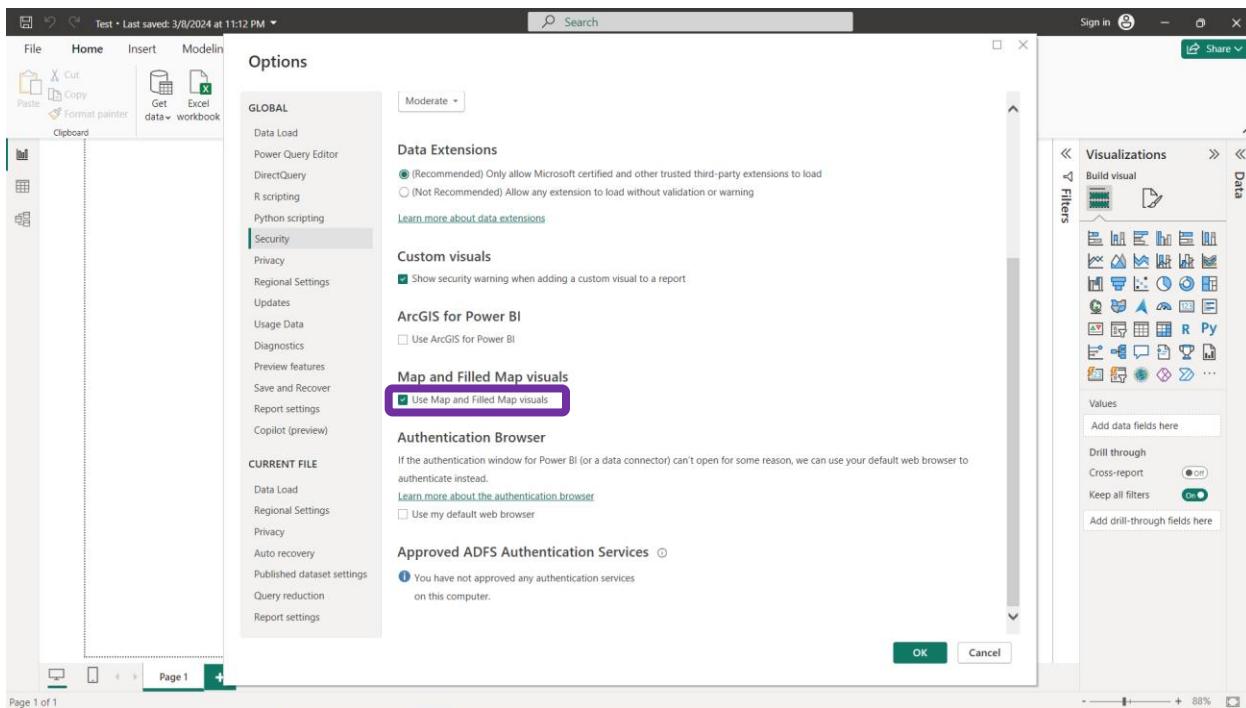
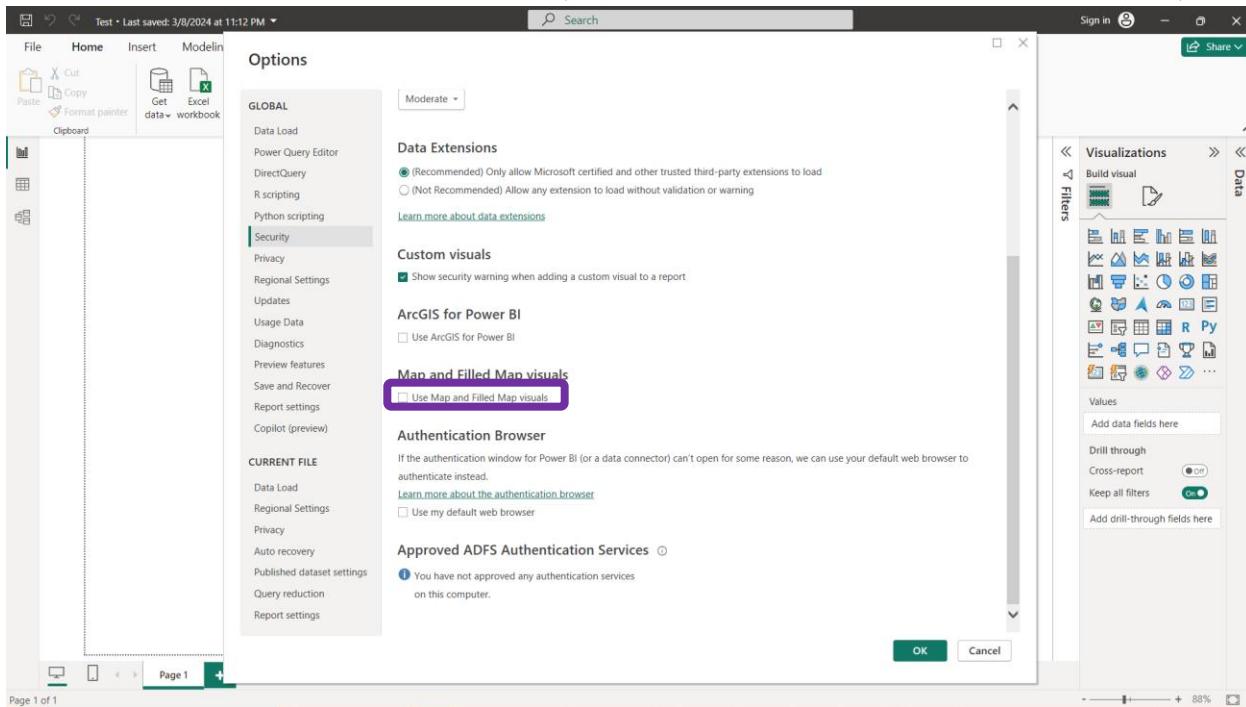




زیر سربرگ Security رو انتخاب میکنیم.



و میریم پایین تا تیک Use Map & Filled Map visuals رو بزنیم.

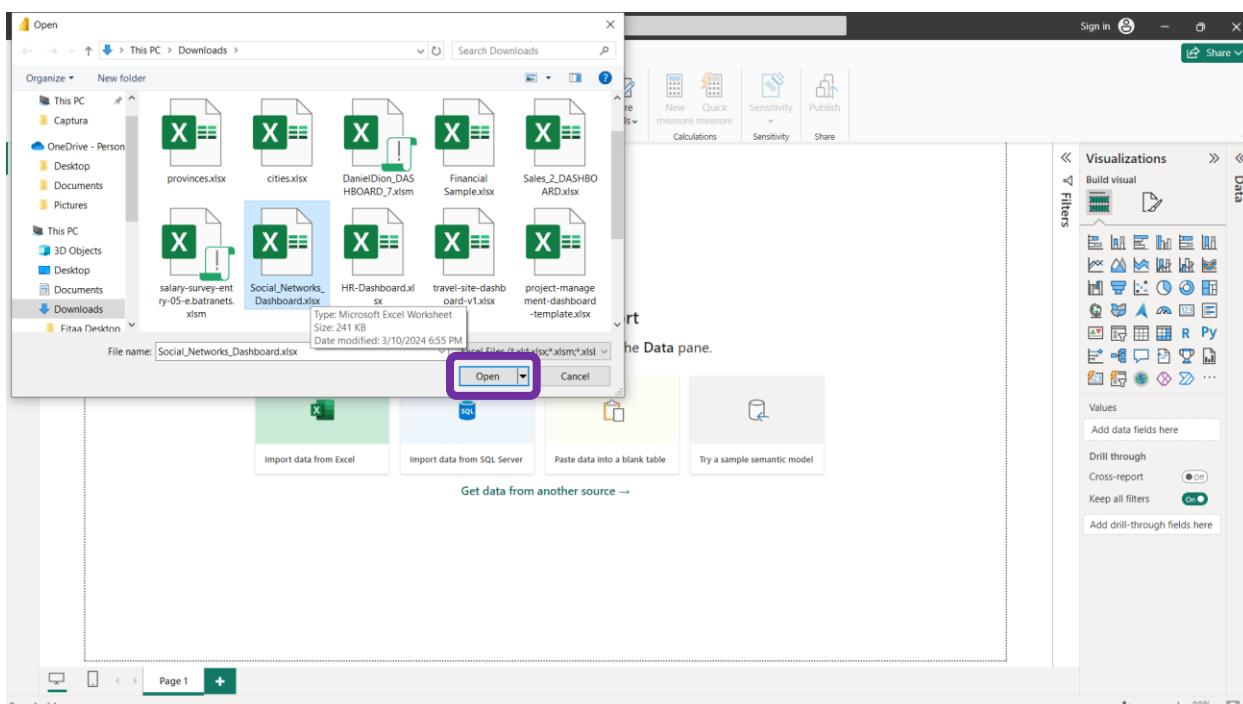
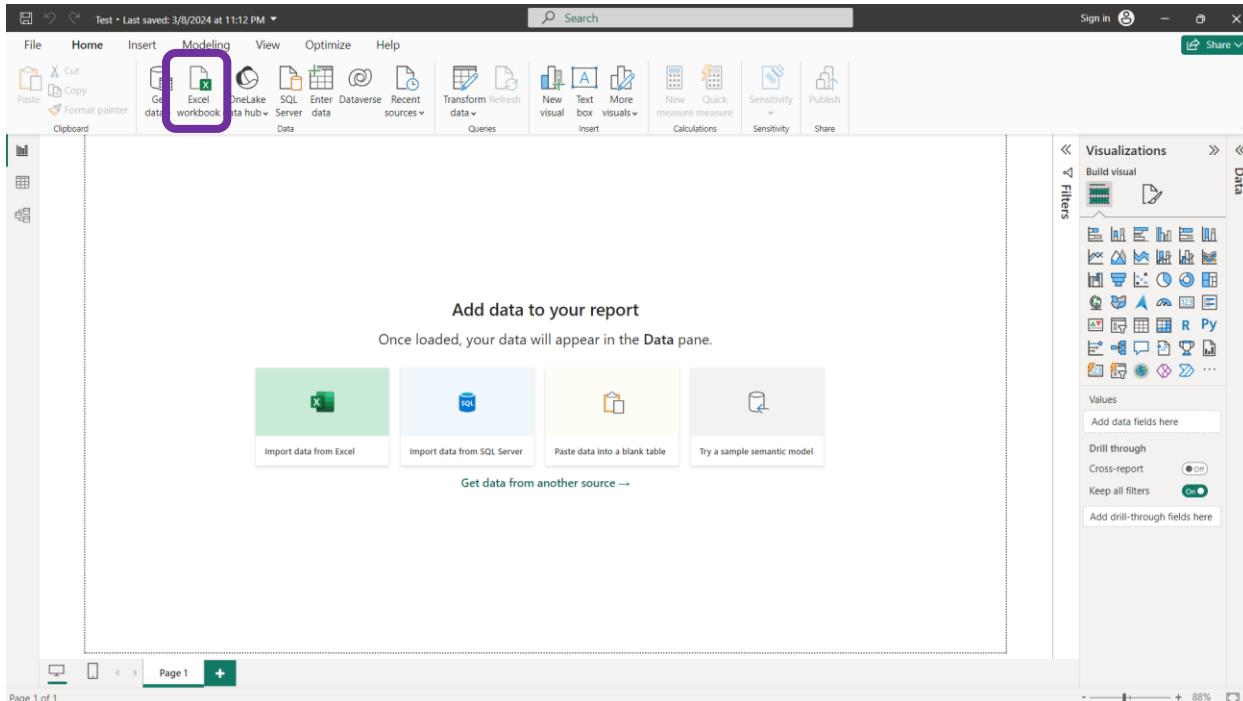


این به صورت پیش فرض غیرفعاله ولی باید فعالش بکنید و OK رو بزنیم.

برای اینکه بتونیم ازش استفاده کنیم باید صفحه رو بیندم و مجدد Power BI رو اجرا میکنیم.

برای اینکه اولین Dashboard مون رو بسازیم به یک فایل Data احتیاج داریم که ما به بررسی شبکه های اجتماعی میپردازیم.

خب توى صفحه ئى Power BI از اون بخش رىبون Excel workbook رو انتخاب مىكتىم و فايل مورد نظرمون رو انتخاب مىكتىم و Open رو ميزنيم.



T_DATA رو میز نیم و Preview را کلیک می کنیم. تیک میز نیم و بعد Load را کلیک می کنیم.

The screenshot shows the Power BI desktop interface. In the center, there's a preview window titled "T_DATA" displaying a table with six columns: MONTH, FB Audience Count, FB Posts, FB Likes, FB Shares, and FB Con. The data is as follows:

MONTH	FB Audience Count	FB Posts	FB Likes	FB Shares	FB Con
1/1/2016	50000	19	69	135	
2/1/2016	51500	20	60	135	
3/1/2016	52000	17	53	134	
4/1/2016	52500	20	60	121	
5/1/2016	52700	17	65	125	
6/1/2016	55000	17	56	146	
7/1/2016	57000	19	66	130	
8/1/2016	62000	17	56	136	

A message at the bottom of the preview window says: "The data in the preview has been truncated due to size limits." At the bottom right of the preview window, there are three buttons: "Load" (highlighted with a purple box), "Transform Data", and "Cancel".

داره فایل ما رو Load میکنه. اطلاعاتش رو بر میداره میاره.

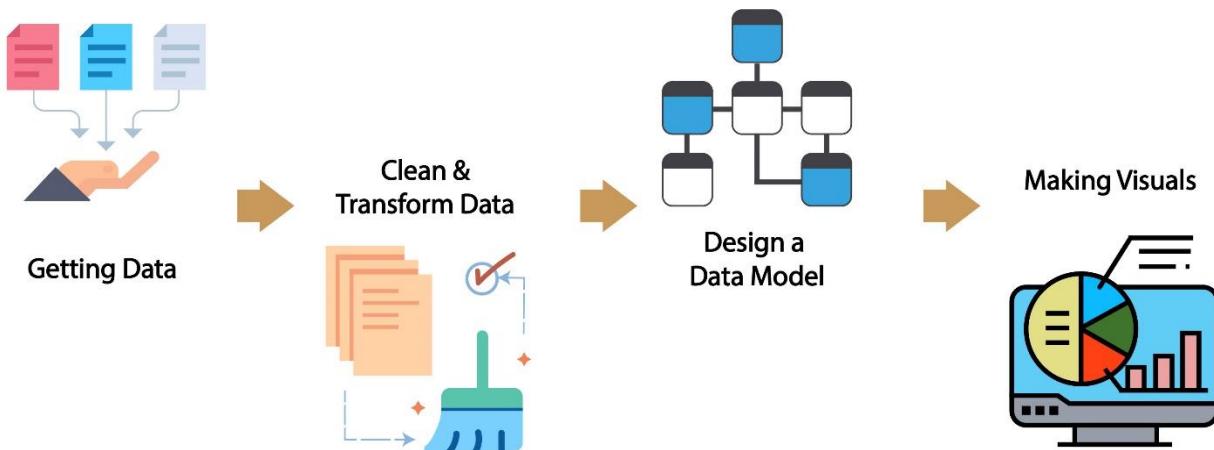
The screenshot shows the Power BI desktop interface again. In the center, a dialog box titled "Add" is open, with "Load" being the selected option. The text inside the dialog says: "Once loaded, you can..." followed by a list item: "T_DATA Creating connection in model...". Below the dialog, there are four buttons: "Import data from Excel", "Import data from SQL Server", "Paste data into a blank table", and "Try a sample semantic model". At the bottom of the dialog, there is a link: "Get data from another source →".

یک منتظر میمونیم Load Social Networks میشه.

پیاده سازی پروژه در Microsoft Power BI

ما برای اینکه بتوانیم توی Power BI گراف ها و ویژوال ها و داشبورد هامونو ایجاد کنیم باید ابتدا مراحل زیر را طی کنیم:

1. اول از همه باید مرحله Get data را انجام بدم یعنی درواقع داده هامونو وارد سیستم بکنیم.
2. و بعد از اون باید داده هامونو تمیز و آماده بکنیم مرحله Clean & Transform مرحله ایه که توی اون ما داده های زائدمونو حذف میکنیم و تبدیل های لازم رو برای اینکه داده قابل استفاده بسه انجام میدیم.
3. مرحله بعدی یعنی Design Data Model یا طراحی مدل داده ارتباط بین داده هارو مشخص میکنیم و جداول داده هارو بهم متصل میکنیم.
4. در مرحله چهارم یا Making Visuals ما در این مرحله گراف ها و داشبورد هامونو ایجاد میکنیم.



مرحله اول: Getting Data

در این مرحله Power BI برای ما ابزار های بسیار گسترده ای را فراهم کرده تا بتوانیم از Source های متنوعی از داده ها استفاده بکنیم از فایل ها ، دیتابیس ها ، ابزار های Power BI ، کلود Azure ، سرویس های آنلاین و ...

لیست مواردی که میتوانیم بهشون متصل بشیم و سرویس هایی که Power BI باهاشون ادغام شده تعدادشون خیلی زیاده و تقریباً میشه مطمئن بود که همیشه میتوانیم به داده ها دسترسی داشته باشیم.



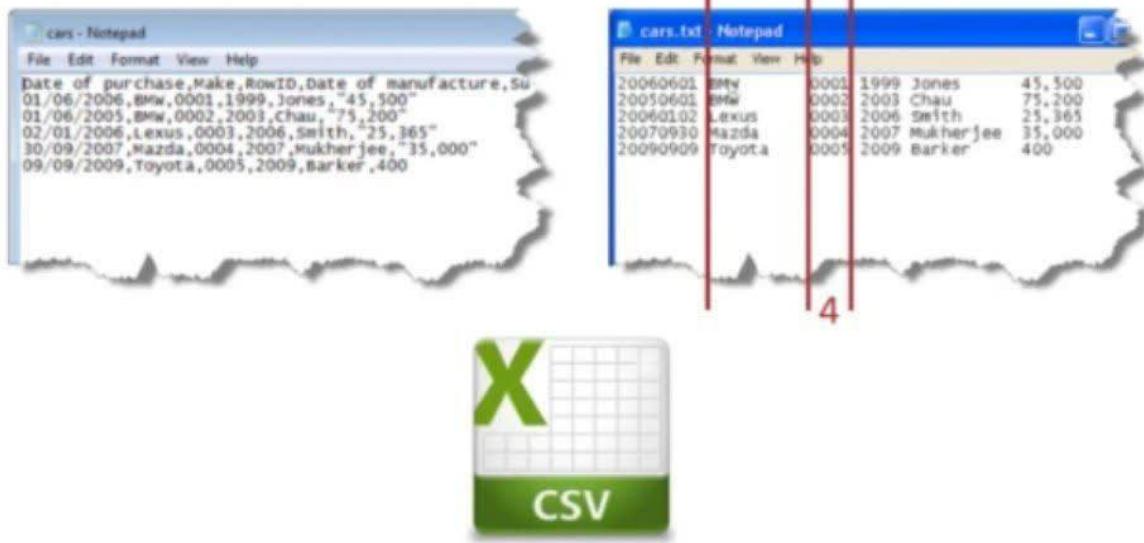
چند نوع از فایل های داده خیلی رایج:

1. فایل CSV که داده ها داخل اون به صورت تکست ذخیره میشن و فیلد های داده شون به دو مدل مختلف از هم دیگه تقسیم میشن:
- 1.1 یک مدل با یک کاراکتر خاص مثل Comma ، Dash ، Tab از هم جدا میشن.

Getting Data

و یک مدل دیگه Fixed Width هست یعنی یک عرض ثابت دارن یعنی طول هر کدام از فیلد ها ثابت هست و با تعداد کاراکتر ها از هم جدا نمیشوند.

CSV File



2. فایل های XML یک مدل دیگه از فایل هایی که داده در آن ذخیره میشوند است و فیلد های داده با یه سری از تگ ها به شکل زیر از هم جدا میشوند.

```
<?xml version="1.0" encoding="UTF-8"?>
- <EmployeeData>
  - <employee id="34594">
    <firstName>Heather</firstName>
    <lastName>Banks</lastName>
    <hireDate>1/19/1998</hireDate>
    <deptCode>BB001</deptCode>
    <salary>72000</salary>
  </employee>
  - <employee id="34593">
    <firstName>Tina</firstName>
    <lastName>Young</lastName>
    <hireDate>4/1/2010</hireDate>
    <deptCode>BB001</deptCode>
    <salary>65000</salary>
  </employee>
</EmployeeData>
```

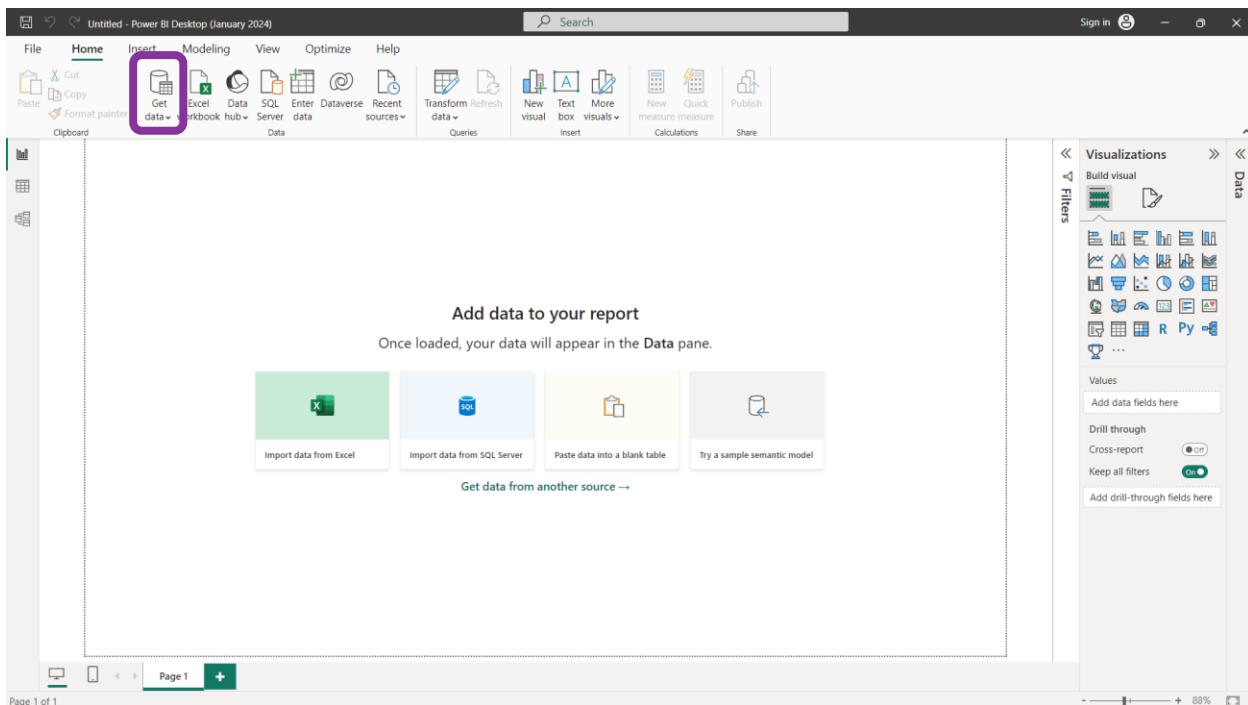
3. یک نوع دیگه از داده داده هایی است که به صورت JSON ذخیره میشوند و در ذخیره سازی در این حالت تفکیک ها با آکلاد و براکت اتفاق میافتد و عنوان ها داخل double quotation قرار میگیره و با دو نقطه نشون میده مقدار اون چی هست.

```

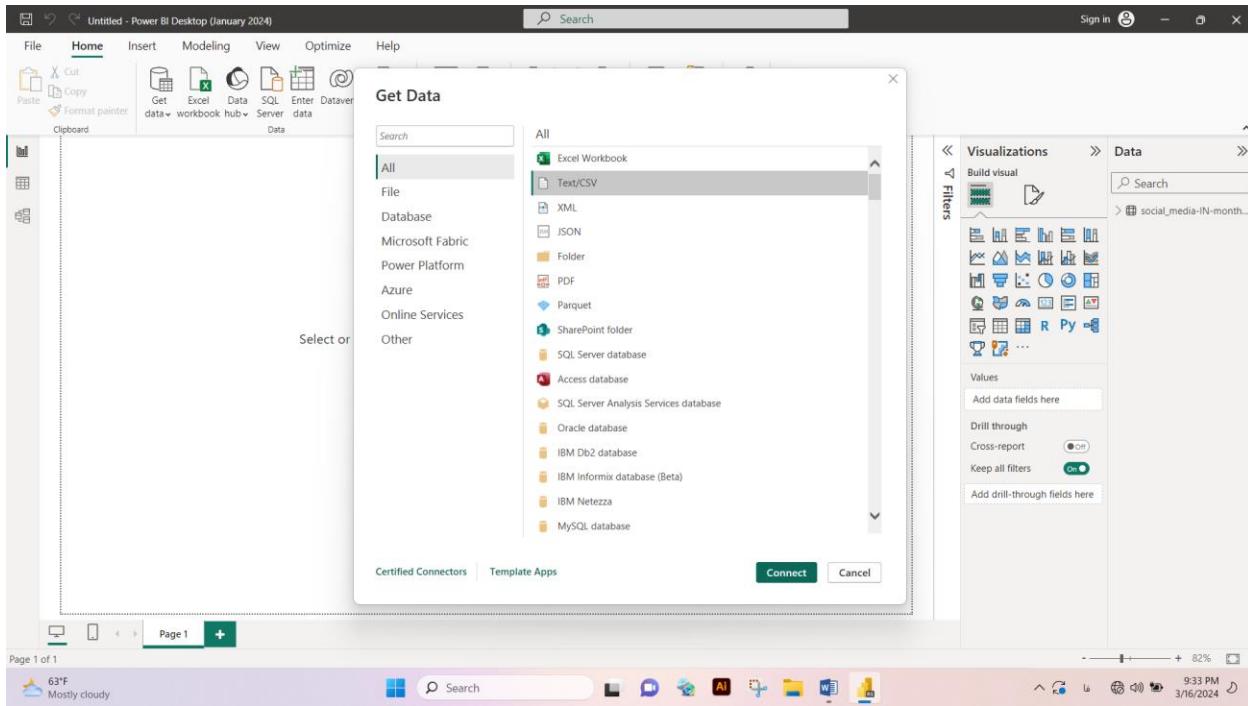
[{"height": 64, "width": 251, "x": 781, "y": 1306}, {"height": 41, "width": 128, "x": 216, "y": 296}]

```

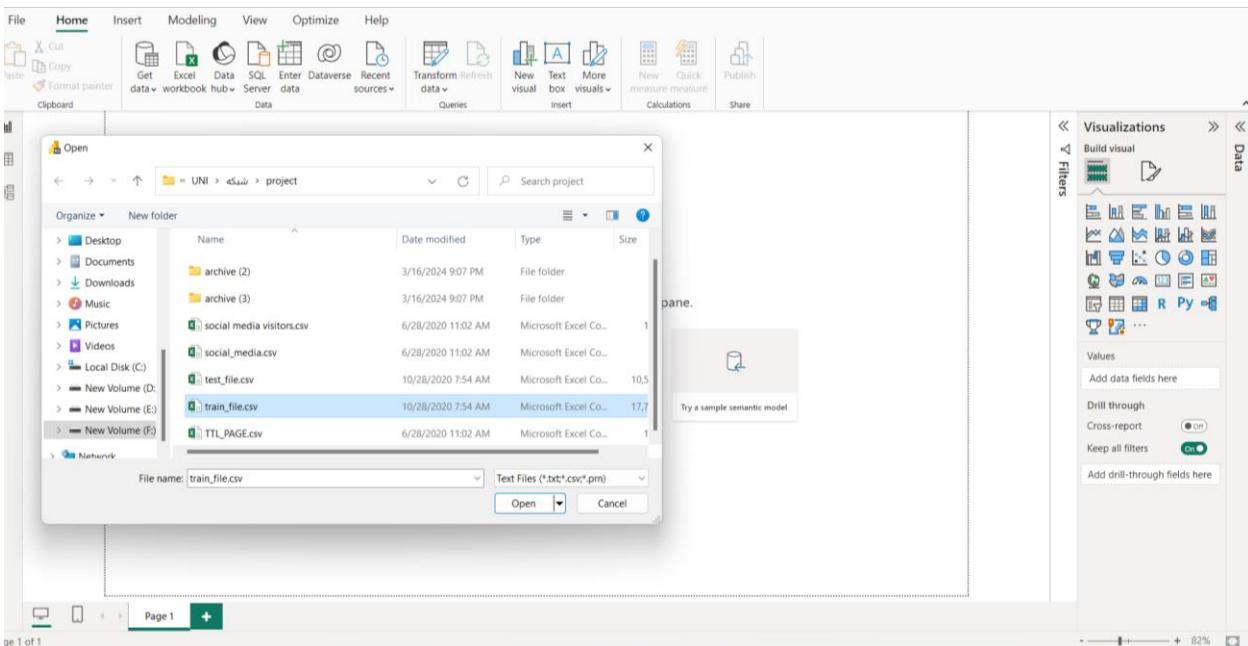
برای دریافت داده در Power BI ما باید از Get Data استفاده کنیم.



پنجره Get Data برای ما نمایش داده میشے. داخل این پنجره انواع سرویس هایی که ما میتوانیم ازشون داده دریافت کنیم لیست شده و ما میتوانیم بسته به اون داده ای که داریم مدلمنو انتخاب کنیم. در این بخش بسته به نوع فایلمنون همونطور که در بالا اشاره شد نوع فایلمنونو انتخاب میکنیم.



حالا در این بخش فایل مورد نظر را انتخاب و **Open** رو میزنیم.



حالا یه پیش نمایش از داده ها به صورت جدول بهمن میده و **Load** رو میزنیم.

حال دیتابست هامون در این قسمت نمایش داده میشه.

حالا الان ما مرحله Get data رو انجام داديم حالا مرحله دوم تمیز و تبدیل کردن داده هاست.

مرحله دوم: Clean & Transform Data

خوب حالا توي اين بخش ميخوايم داده هامون رو تميز و تبديل کنيم.
حالا ميخوايم با يکي از ابزار هاي خيلي جذاب Microsoft به نام Power Query آشنا بشيم.

این ابزار دقیقاً توي مرحله Get data وقتی که داده ها دارند وارد میشنوند وارد عمل میشود.

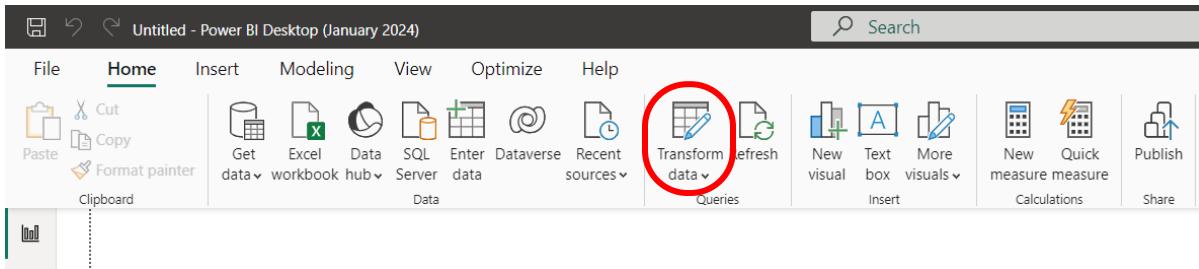
برای وارد شدن به این قسمت دوتا راه وجود داره: یکیش وقتی که میخوايم دیتا رو لود کنیم در همین قسمت میتونیم گزینه Transform Data رو بزنیم.

Clean & Transform Data



The screenshot shows the Microsoft Power BI desktop application. In the center, there's a preview of a CSV file named "dummy_data.csv". The "Transform Data" button, which is green and outlined in red, is located at the bottom right of the preview window. The ribbon menu at the top includes "File", "Home", "Insert", "Modeling", "View", "Optimize", and "Help". On the right side, there's a "Visualizations" pane and a "Data" pane. The "Data" pane contains a search bar and a list of available data sources.

يا راه ديگه اينکه میتونیم از ریبون بالای Power BI وارد محیط Power Query بشیم.



با زدن **Transform Data** وارد **Power Query Editor** میشیم.

	Age	Age range	Gender	time_spent	Platform	Interests
1	36	over 55	male		Instagram	Sports
2	46	41-55	female		Facebook	Travel
3	32	26-40	male		Instagram	Sports
4	60	over 55	non-binary		Instagram	Travel
5	25	18-25	male		Instagram	Lifestyle
6	38	26-40	male		Facebook	Travel
7	56	over 55	male		YouTube	Sports
8	36	26-40	male		Instagram	Sports
9	40	26-40	non-binary		YouTube	Lifestyle
10	28	26-40	non-binary		Instagram	Sports
11	28	26-40	female		YouTube	Lifestyle
12	41	41-55	non-binary		Facebook	Sports
13	53	41-55	non-binary		Facebook	Sports
14	57	over 55	female		Instagram	Lifestyle
15	41	41-55	non-binary		YouTube	Travel
16	20	18-25	male		Facebook	Sports
17	39	26-40	male		Instagram	Travel
18	19	18-25	female		Instagram	Sports
19	41	41-55	female		Facebook	Sports
20	61	over 55	female		Facebook	Sports
21						

در سمت چپ لیست جداولی هست که وارد **Power BI** کردیم.

کلیه مراحلی که قراره انجام بدیم در بخش **Applied Steps** قراره یک به یک ذخیره بشه.

این فایل ، فایلی است که بر اساس **Age** ، **Gender** ، **Age range** ، **Platform** ، **Time spent** ، **Interests** ، **Location** ، **demographics** ، **Profession** ، **Income** ، **Income range** ، **Home Owner** ، **Indebt** ، **Owns car** دسته بندی شده است.

اگه ما بخوایم اسامی برنامه های social media مون رو به عنوان اسم ستون ها قرار بدیم. این کار رو میتوانیم با کلیک روی بخشی که در پایین مشخص کردیم و بعد زدن **Use First Row as Header** انجام بدیم. وقتی این رو بزنیم میداد اتوماتیک هر آنچه که در ردیف اول بود رو به عنوان اسم ستون ها برای ما قرار میده. که البته Power Query الان اینطوری هست و نیازی به تغییر نیست.

The screenshot shows the Power Query Editor interface. At the top, the ribbon has tabs like File, Home, Transform, Add Column, View, Tools, and Help. Below the ribbon, there's a toolbar with icons for Close & Apply, New Source, Recent Sources, Enter Data, Data source settings, Manage Parameters, Refresh Preview, Properties, Advanced Editor, and various column management tools. The main area shows a table with columns: age, gender, and time_spent. A context menu is open over the first row, with the option 'Use First Row as Headers' highlighted by a purple oval.

مثلا الان ما Use First Row as Header رو زديم و الان سطر اول رو به عنوان اسم ستون ها در نظر گرفته.

This screenshot shows the Power Query Editor after applying the 'Use First Row as Headers' transformation. The table now has five columns: male, Instagram, Sports, United Kingdom, and Urban. The 'Properties' pane on the right shows the applied step 'Changed Type1' under 'APPLIED STEPS'. The status bar at the bottom provides information about the data: 12 COLUMNS, 999 ROWS, and a download timestamp of 3/18/2024 at 10:51 PM. The system status bar at the bottom left shows the temperature as 56°F and the weather as mostly cloudy.

اگه بخوايم اسم ستون و عوض کنيم دوبار روش کليک ميکنيم و به صورت دستي اسم ستون رو تغيير ميديم.

The screenshot shows the Microsoft Power Query Editor interface. A table named 'dummy_data' is displayed with 28 rows of data. The columns are: age, gender, time_spent, platform, interests, location, and demographi. A context menu is open over the 'age' column, specifically on the 'Change Type' option under the 'Transform' tab. The 'APPLIED STEPS' pane on the right shows a step labeled 'Changed Type'. The status bar at the bottom indicates 'PREVIEW DOWNLOADED AT 10:33 PM'.

يك نكته‌اي که باید بهش توجه بکنیم نوع داده هستش که وقتی وارد Power BI می‌کنیم مثلًاً نوع این داده عددی هستش حالا اگه بخوايم نوع داده رو عوض کنیم روی آیکون اون کلیک می‌کنیم و می‌بینیم که می‌توانیم انواع مختلف داده را برای این ستون انتخاب کنیم.

The screenshot shows the Microsoft Power BI - Power Query Editor interface. A table named 'Data' is displayed with 28 rows of data. The columns are: Age, Age range, Gender, time_spent, platform, interests, and location. The 'Age' column has been converted to 'Age range' as shown in the 'APPLIED STEPS' pane. The status bar at the bottom indicates 'PREVIEW DOWNLOADED AT 10:15 PM'.

يه راه ديگه‌اي که وجود داره راست کلیک روی اون ستون هستش اگه روی Change Type کلیک کنیم می‌توانیم مختلف داده رو واسه اين داده انتخاب کنیم حالا اين تایپ درست هستش و تغييرش نميديم.

اگه ردیف اضافه داشته باشیم و بخوایم با یک الگوی استاندارد حذف کنیم میتوانیم از بخش Remove Rows میتوانیم تعداد سطرهایی که میخوایم حذف بشه رو وارد کنیم.

اگه ما یه سری مقادیر و محتوای خاص رو با یه محتوای دیگه عوض کنیم و اسه اینکار کافیه ابتدا ستون مورد نظر رو انتخاب و بعد Replace Values رو بزنیم و در جای خالی اول چیزی که میخوایم عوض شه رو وارد میکنیم و در جای خالی بعد چیزی که میخوایم جاش بیار رو وارد میکنیم.

This screenshot shows the Power BI Data Editor interface. A 'Replace Values' dialog box is open, overlaid on a table. The table has columns: Age, Age range, Gender, time_spent, Platform, Interests, and Location. The 'Gender' column currently contains 'non-binary' values. The 'Replace Values' dialog box has 'non-binary' in the 'Value To Find' field and 'uncharted' in the 'Replace With' field. The 'OK' button is visible at the bottom right of the dialog.

و میبینیم که نتیجه به این صورت درمی‌آید. و بعد برای اینکه تغییراتمون اعمال بشه Close & Apply رو میزیم. و بعد به صفحه اصلی Power BI برمیگردیم.

This screenshot shows the Power BI Data Editor after applying the changes. The 'Close & Apply' button in the top-left corner is circled in red. An orange arrow points from the 'Replace Values' dialog in the previous screenshot to this button. The table now shows 'uncharted' instead of 'non-binary' in the Gender column. The 'APPLIED STEPS' pane on the right shows the 'Replaced Value' step.

و ما میتوانیم توی بخش دیتا جدول مرتبی که ایجاد کردیم رو بینیم.

The screenshot shows the Microsoft Power BI Desktop interface. The top menu bar includes File, Home, Help, and Table tools. The Table tools ribbon has several options: Name, Data, Mark as date table, Manage relationships, New, Quick, New measure column, New table, and Calculations. Below the ribbon is a search bar and a sign-in button. The main area displays a data grid with 1,000 rows of data. The columns include: range, Gender, time_spent, Platform, Interests, Location, Place of living, Profession, Income, Income range, Indebt, isHomeOwner, and Owns_Car. The data consists of various demographic and behavioral information. At the bottom of the screen, there is a taskbar with icons for various applications like File Explorer, Edge, and Mail, along with system status indicators like battery level and signal strength.

برای اینکه کیفیت داده هامون رو ارزیابی کنیم میتوانیم دوباره وارد بخش Transform Data بشیم و تیک قسمت Column Quality را فعال کنیم. و میبینیم که داده خالی نداریم.

The screenshot shows the Microsoft Power Query Editor interface. The top menu bar includes File, Home, Transform, Add Column, View, Tools, and Help. The ribbon has sections for Formula Bar, Monospaced, Column distribution, Always allow, Show whitespace, Column profile, Go to Columns, Advanced Editor, Dependencies, and Promoted Headers. The left sidebar shows 'Query Settings' with 'Name' set to 'Data' and 'All Properties'. The right sidebar shows 'APPLIED STEPS' with 'Source', 'Navigation', 'Promoted Headers', and 'Changed Type' selected. The main area displays a data preview table with columns: Age, Gender, time_spent, Platform, Interests, and Location. Each column has a 'Column quality' section showing the percentage of valid, error, and empty values. The preview table contains 1,000 rows of data. The bottom of the screen shows a taskbar and system status indicators.

و همچنین Column distribution رو اگر فعال کنیم میبینیم که یک نمودار کوچک بر حسب Unique بودن و Distinct بودن برای ما فعال میشود.

Power BI Power Query Editor showing a transformed dataset. The 'APPLIED STEPS' pane shows a 'Changed Type' step applied to the 'time_spent' column. The data preview shows various columns like Age, Gender, time_spent, Platform, etc., with their respective data types and validation status.

حالا دوباره باید برای اینکه تغییراتمون اعمال بشے Close & Apply رو بزنیم و بعد به صفحه اصلی Power BI بر میگریدم.

مرحله سوم: Design a Model

Power BI Power Query Editor showing a transformed dataset. The 'APPLIED STEPS' pane shows a 'Changed Type' step applied to the 'time_spent' column. A context menu is open over the 'time_spent' column, specifically on the 'Remove Duplicates' option.

حالا اگه بريم به بخش داده هامون بخش Age Range ميбинيم که 4 سنی داريم که ميتوnim واسشون categories تعریف کنیم و اینارو به شکل elderly, middle-aged, Young, old دسته بندی کنیم.

Table: Data (1,000 rows) Column: Age range (4 distinct values)

55°F Partly sunny 5:18 PM 3/21/2024

خوب ما باید واسه این یه جدول درست کنیم. واسه اینکارابتدا Enter Data رو میزنیم.

Table: Data (1,000 rows) Column: Age range (4 distinct values)

53°F Partly sunny 5:32 PM 3/21/2024

خوب حالا تو این ستون دقیقا همون رنج های سنی که داریم و اسم ستون تو جدول اصلیمون رو وارد میکنیم. تو قسمت Name مینویسیم. Age categories هم

Table: Data (1,000 rows) Column: Age range (4 distinct values)

Name: Age categories

Load Edit Cancel

Data

- Age range
- Gender
- time_spent
- Platform

و همینطور اسم categories هامونو وارد میکنیم و در آخر Load را میزنیم.

Table: Data (1,000 rows) Column: Age range (4 distinct values)

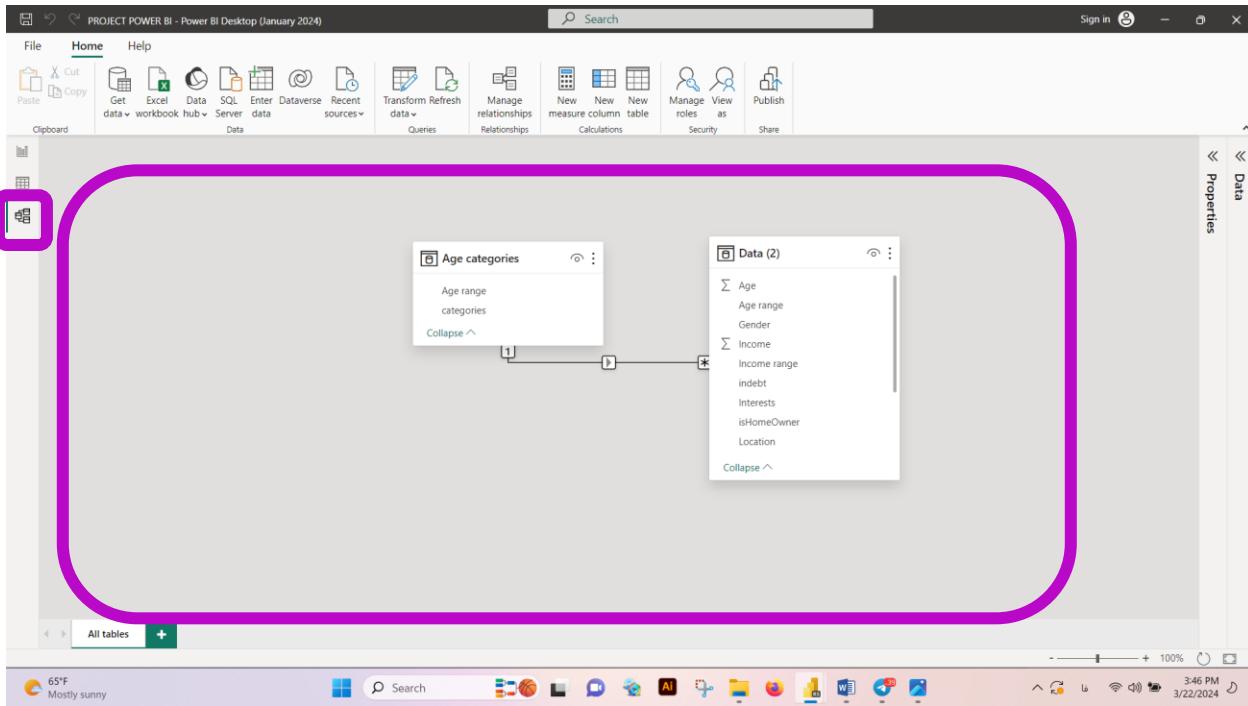
Name: Age categories

Load Edit Cancel

Data

- Age range
- Gender
- time_spent
- Platform

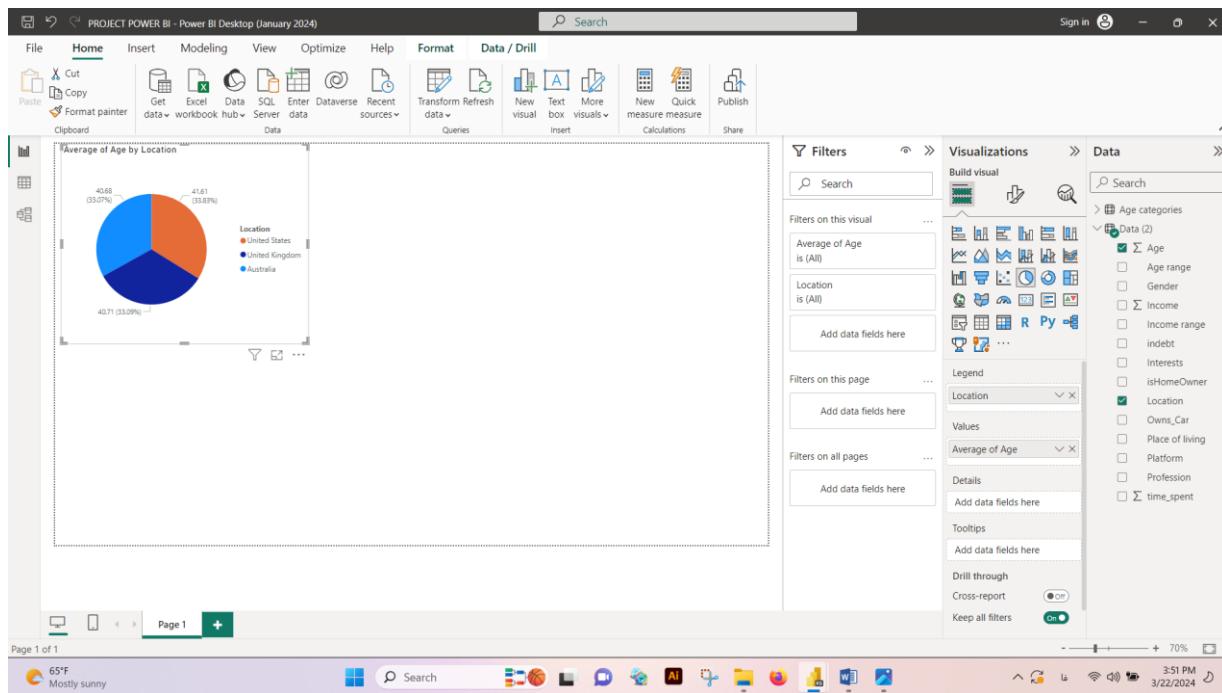
حالا وارد صفحه اصلی Power BI میشیم و به قسمت Relation Model میریم تا ببینیم هامون به چه صورت تعریف شدن.



و الان میبینیم که Relation رو درست شناخته حالا وارد بخش Report View میشیم و میخوایم ویژوال بسازیم.

مرحله چهارم: Making Visuals

در سمت راست ویژوال مورد نظرمونو از قسمت Visualization انتخاب میکنیم و از قسمت Data دیتا ست مورد نظرمونو باز میکنیم و دیتا های مورد نظر با شرایطی که میخوایم بسنجدیم شون رو کلیک میکنیم اگه در جای درست قرار نگرفته بودن Drag & Drop میکنیم و به مکان درستش میبریم.



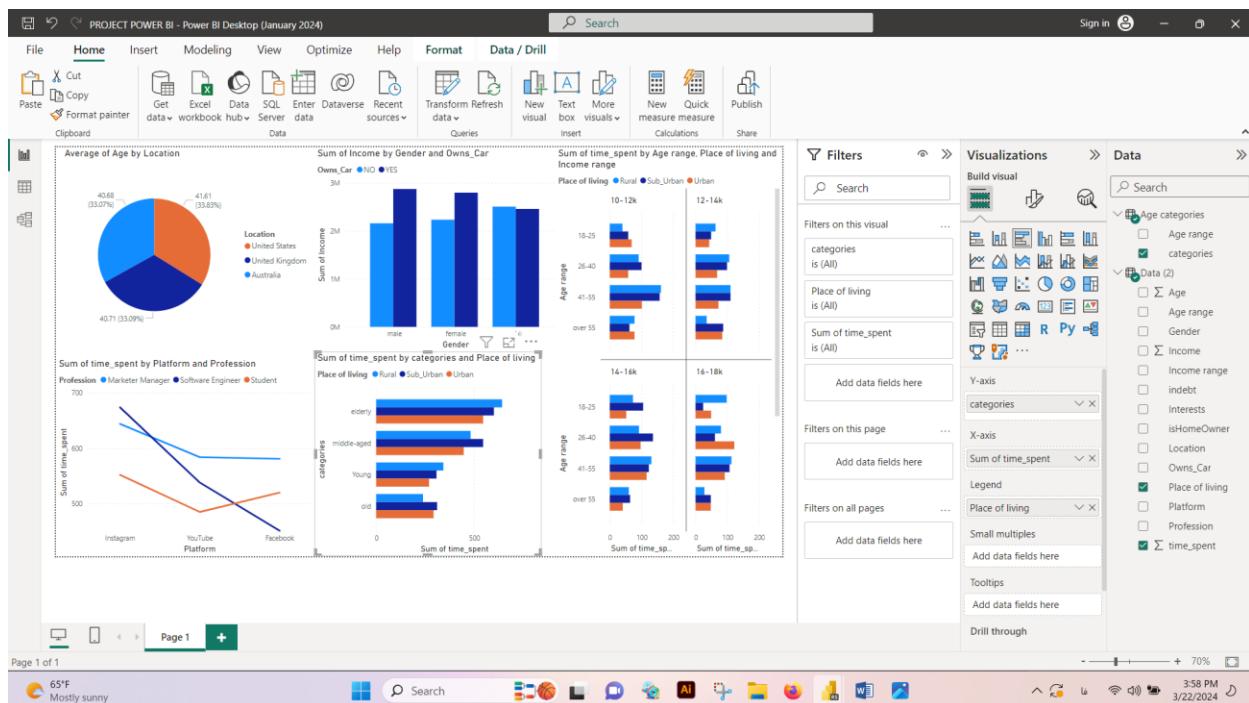
در شکل زیر در نمودار دایره‌ای مون که در واقع اسمش Pie Chart هستش ما بر اساس location و Age Average هستش

این نمودار رسم کردیم

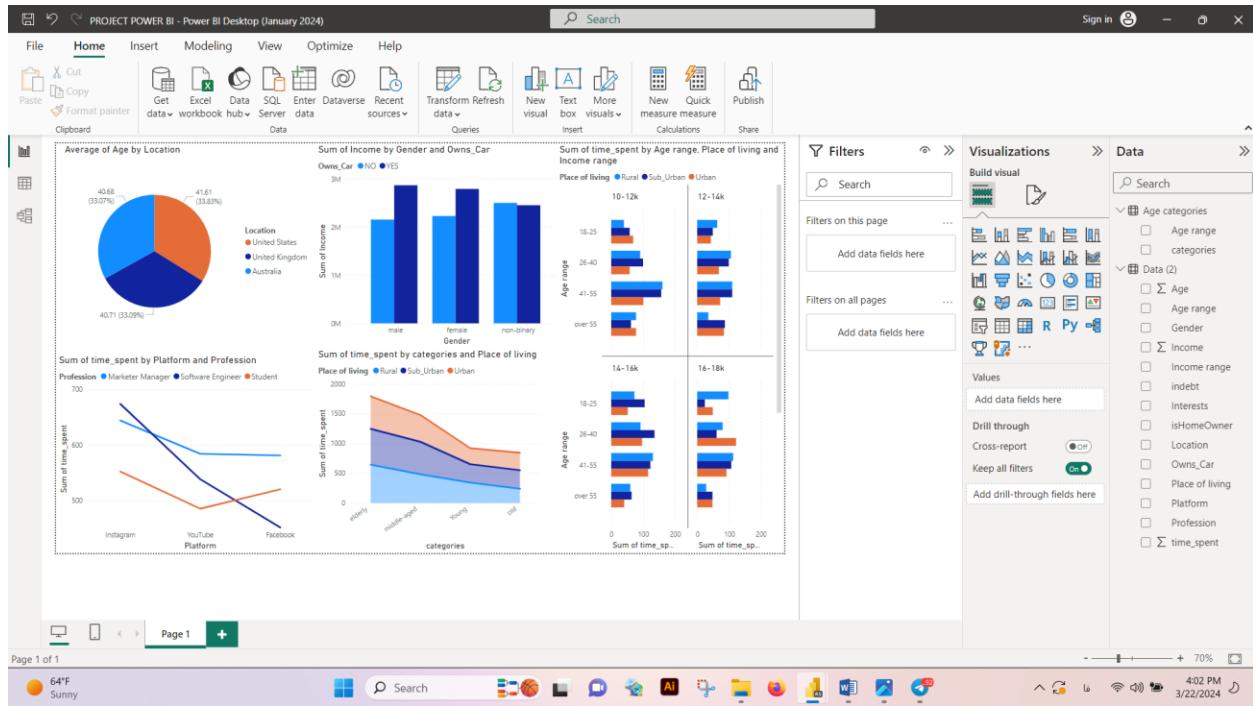
در نمودار بعدی که نمودار ستونی هستش یا Clustered Column Chart بر اساس Gender و Owns-car و Income دسته‌بندی شدن مثلاً مردایی که ماشین از خودشون ندارن در آمدشون حدود ۲ میلیونه و مردایی که ماشین از خودشون دارند حدود ۳ میلیونه این رقم در زنا هم تقریباً مشابه هستش.

خوب تو نمودار سمت چپ پایین که می‌بینیم نمودار خطی هستش اسمش نمودار Line Chart هست و می‌بینیم که بر اساس جمع مدت زمانی که در پلتفرم‌هایی که مد نظر منه یعنی یوتیوب فیسبوک و اینستاگرام می‌گذروند و بر اساس حرفشون تقسیم بندی کردیم.

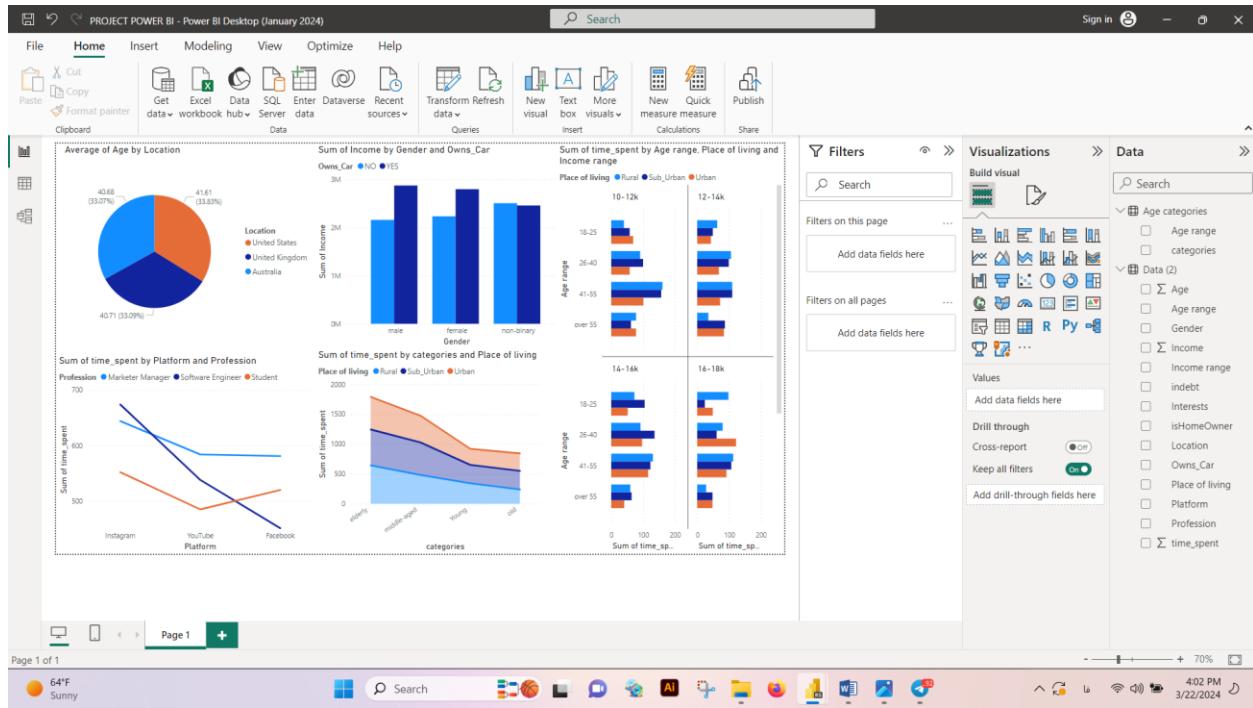
نموداری که در قسمت پایین وسط قرار داره اسمش نمودار Stacked area chart هست و در اون ما بر اساس شهری نیمه شهری یا روستایی بودن زمانی که در پلتفرم‌ها می‌گذروند و محدوده سنیشون رسم شده.



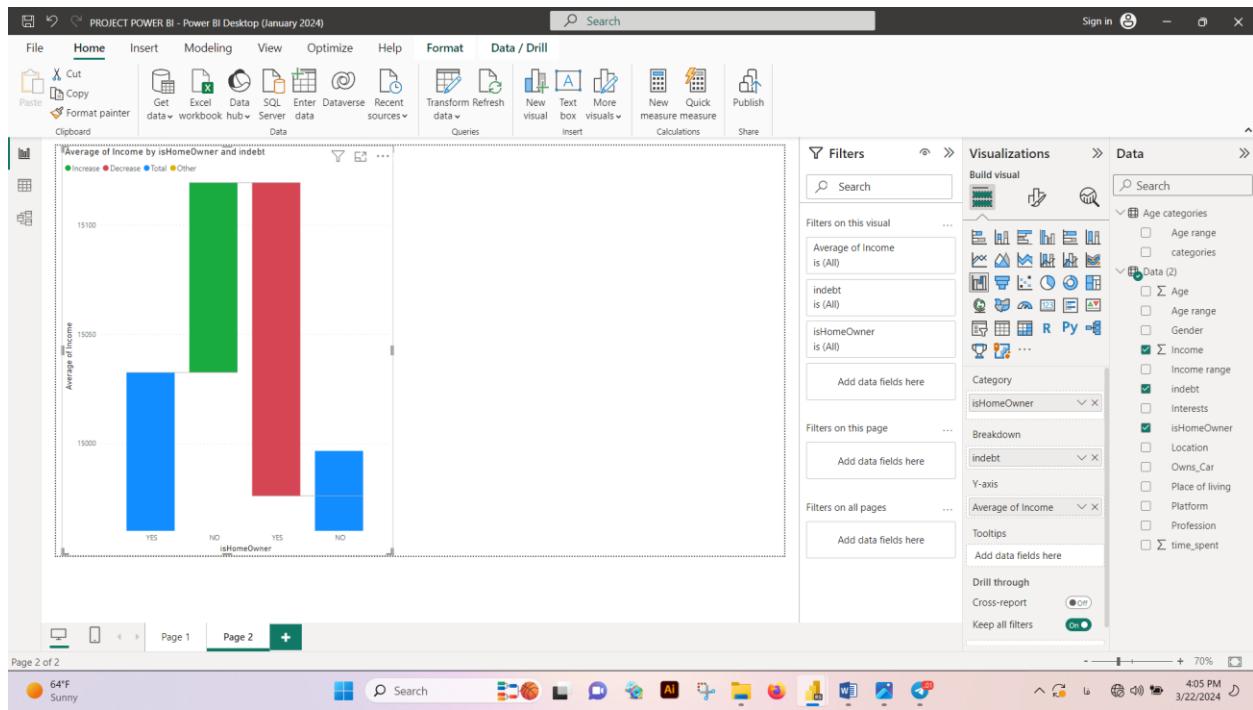
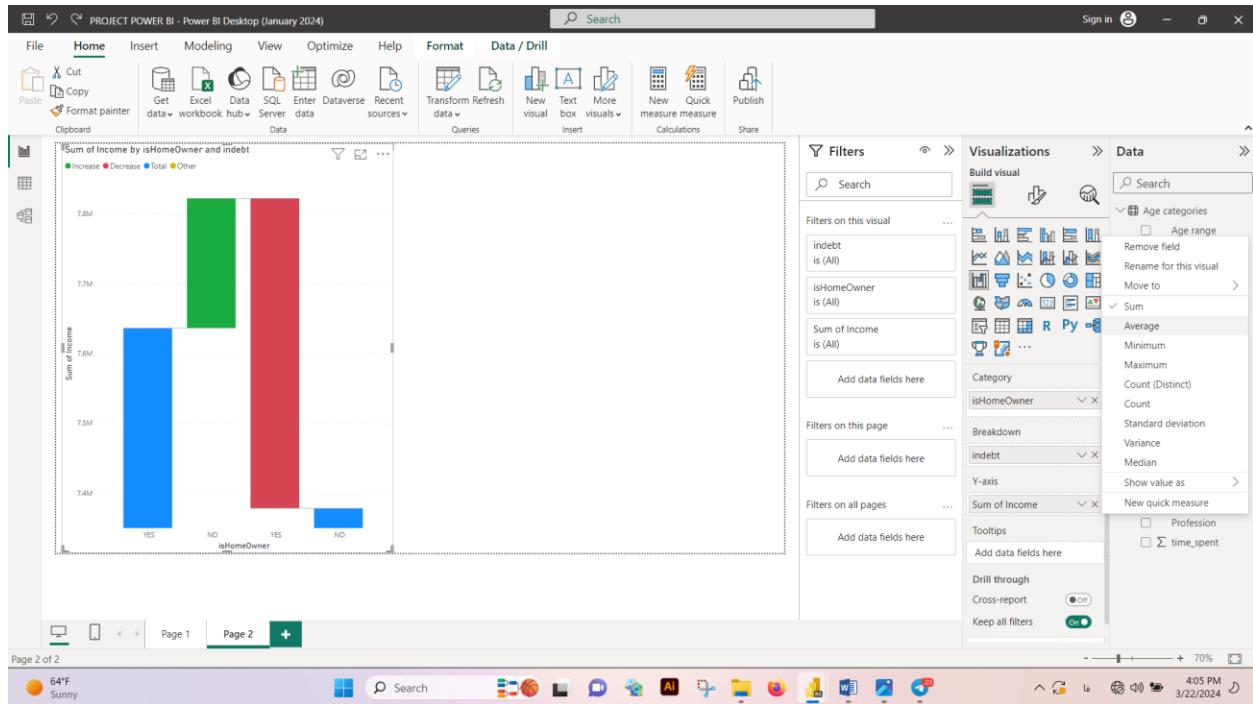
نمودار سمت راست بر اساس جمع مدت زمان و محدوده سنی و مکان زندگی و محدوده درآمد این نمودار مقایسه می‌شه و محدوده درآمدی توی بالای هر نمودار نوشته شده و در سمت چپ هر نمودار محدوده سنی نوشته شده و محدوده زمانی که توی پلتفرم‌های مختلف صرف می‌کنند از قسمت پایین با عدد مشخص شده و همچنین در مکان زندگی با رنگ‌ها مشخص شده.



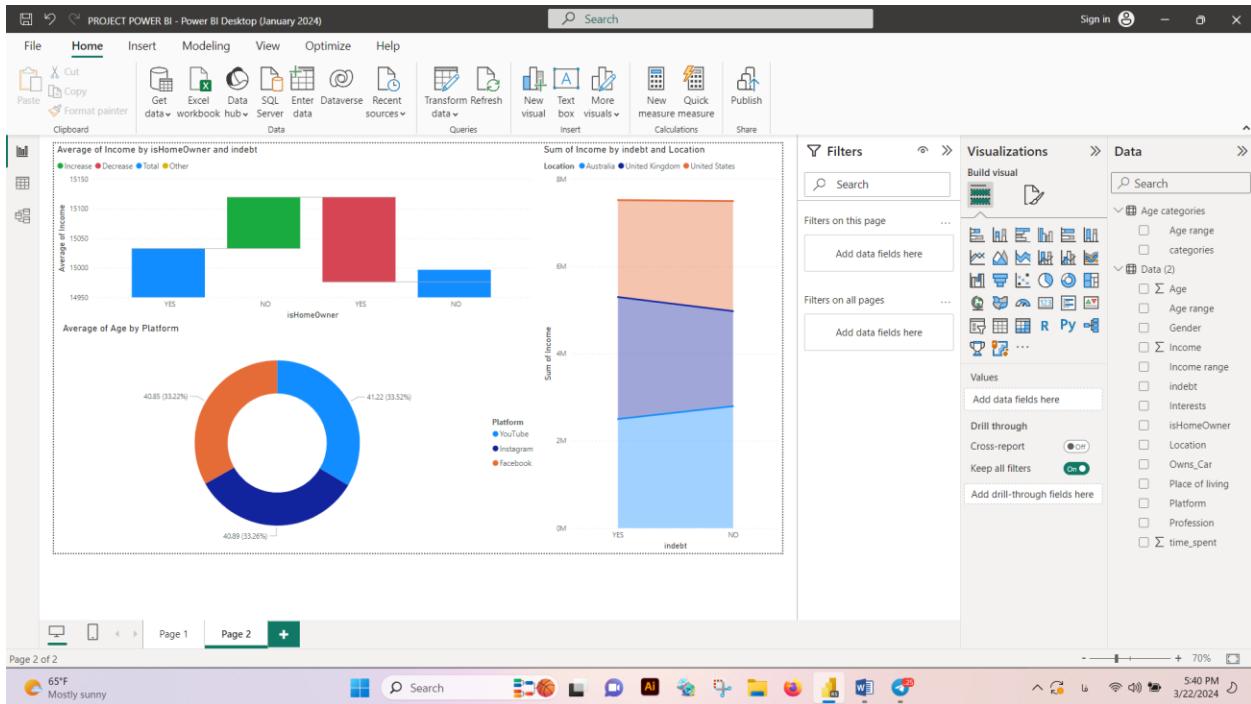
در قسمت پایین میتوانیم صفحه جدیدی ایجاد کنیم و نمودار های بیشتری رسم کنیم.



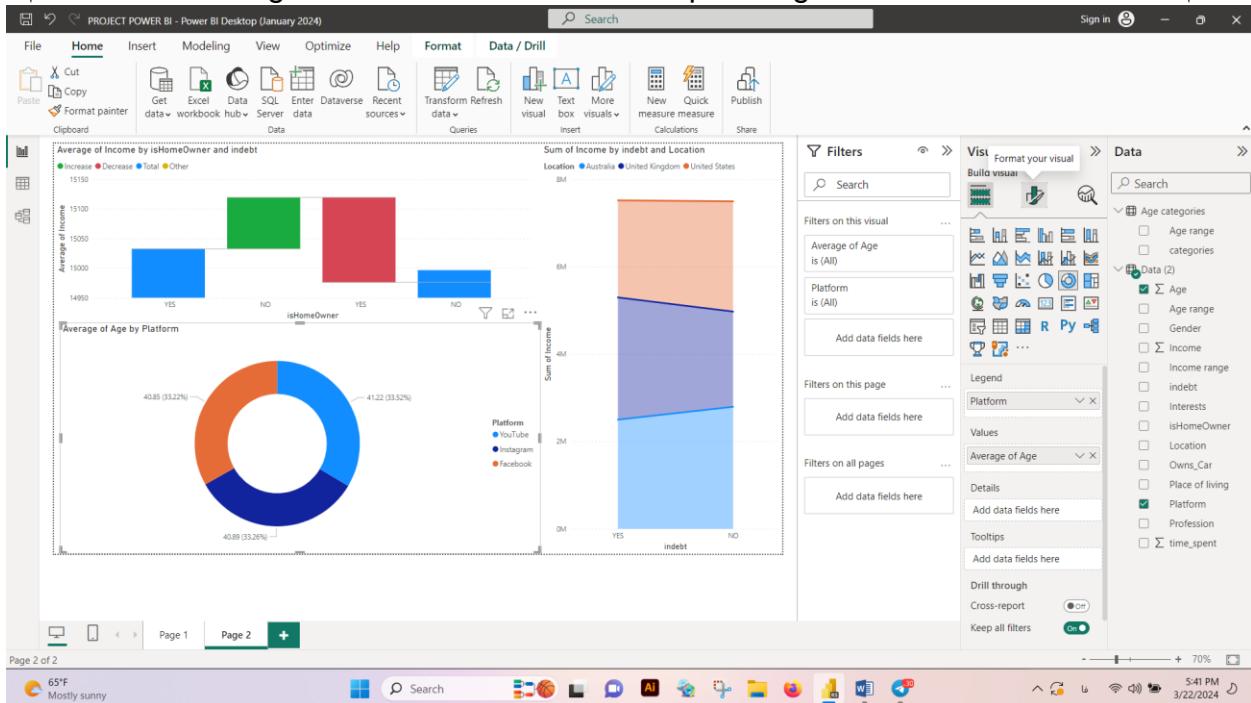
در اینجا نموداری به اسم Waterfall Chart داریم و توی اون نمودارا بر حسب اینکه مالک خونه هستند یا نه و درآمد ها و بدهی های آنها رسم شده است. میتوانیم ایز اون هارو تغییر بدیم.



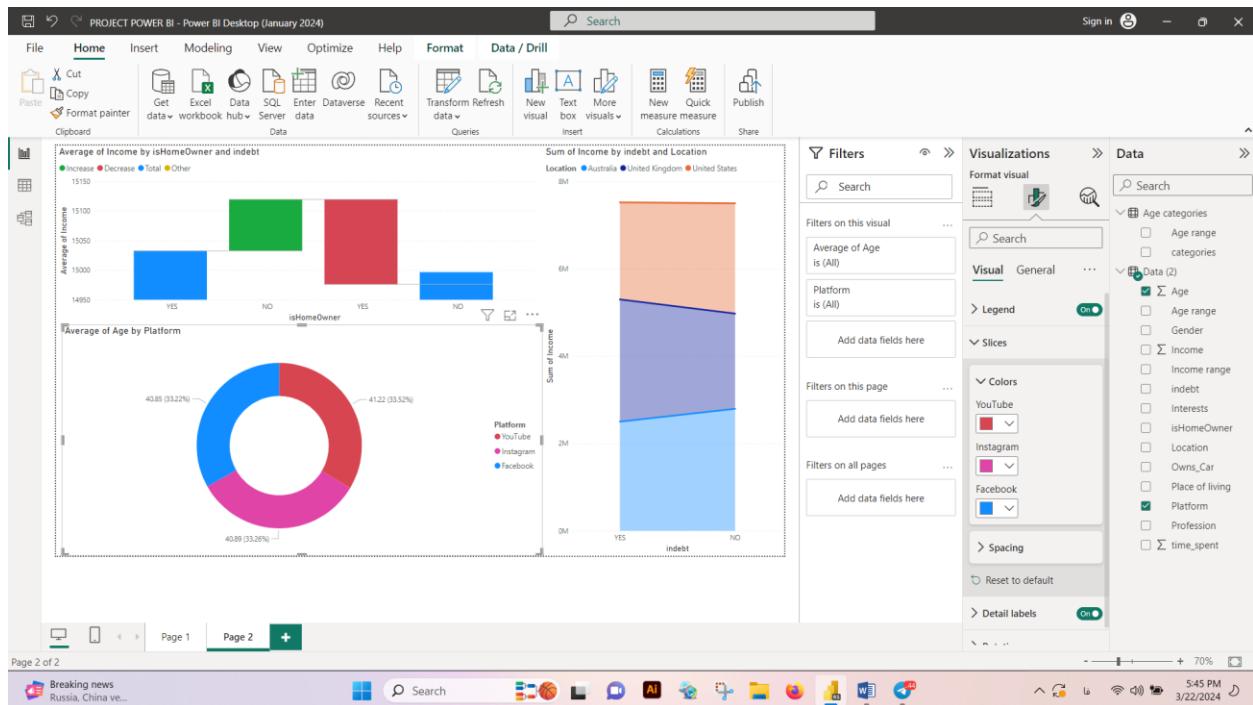
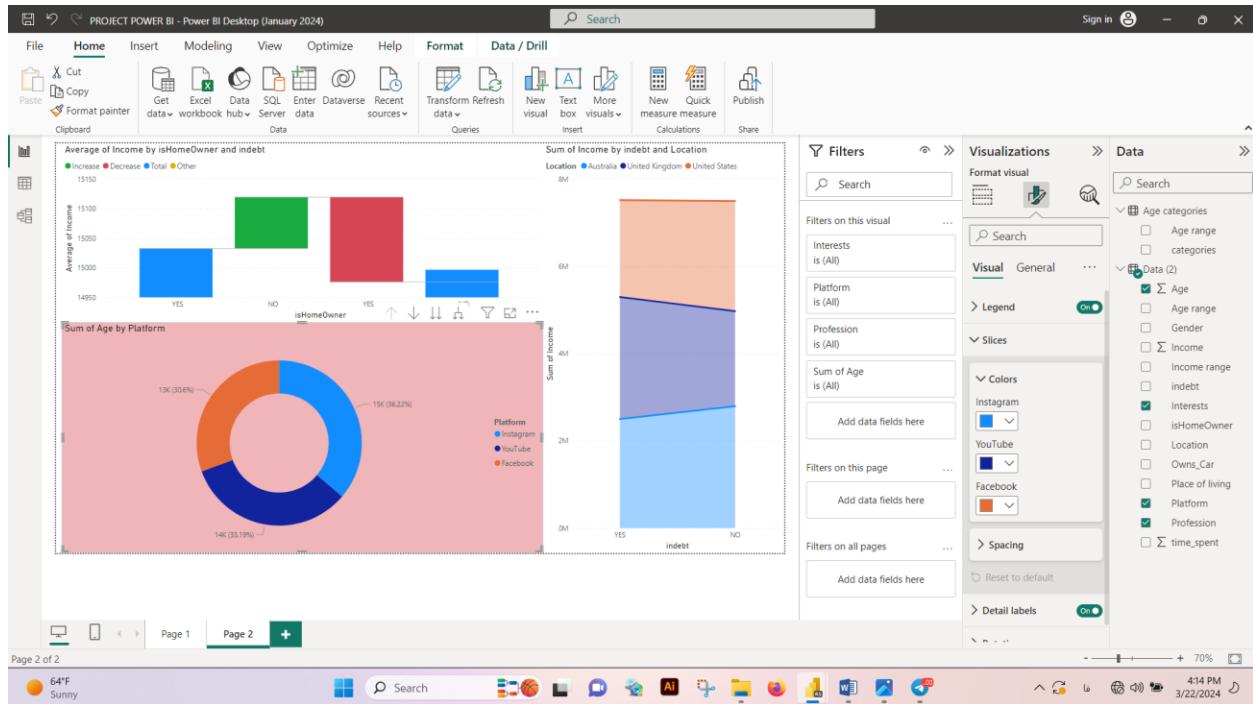
در این قسمت نمودار Donut Chart را داریم که بر اساس پلتفرم ها و میانگین سنی اون ها تقسیم بندی شده.



میتوانیم از بخش Visualization تغییرات توی داشبوردمون بدیم.



برای مثال از قسمت Visuals میتوانیم رنگ slice هارو تغییر بدیم. با خاموش یا روشن کردن میتوانیم اسامی که کنار نمودار ها فعال اند رو فعال یا غیرفعال کنیم.



در قسمت General هم میتوانیم رنگ پس زمینه و نوع فونت و بقیه موارد را تغییر بدیم.

در نمودار سمت راست هم میتوانیم نموداری ببینیم بر اساس کشورها و دارایی‌ها و بدهی‌ها تقسیم‌بندی شده است.

Microsoft Fabric

حالا میخوایم به بررسی و طراحی داشبورد در نرم افزار Microsoft Fabric بپردازیم.

چیست؟ Microsoft Fabric

خب در ابتدا به معرفی Microsoft Fabric می پردازیم. Microsoft Fabric به خاطر این اومده که مشکلات معمول صنعت را حل کند و فابریک یک فناوری تحول آفرین است.

از طریق مطالعه موردنی، مشکلاتی را که Fabric برای حل آنها ساخته شده است، بررسی خواهیم کرد و با چند مفهوم فاندامنتال که باید از آن بدانیم آشنا خواهیم شد:

۱. OneLake
۲. The seven Fabric experiences
۳. The four compute engines of Fabric

نکته مهم در مورد Fabric این است که یک راه حل یکپارچه انتها به انتها است. به این معنی که برای هر شخصیت داده در سازمان شما پاسخگو است. بنابراین اگر در حال حاضر در هر یک از این نقش‌ها:

۱. Data analyst
۲. Power BI developer
۳. Data engineer
۴. Data scientist
۵. Database administrator
۶. Solution architect
۷. Manager/Director

شاید خارج از فابریک، کار می کنید، این ارائه قطعاً برای شما بسیار مرتبط و جالب خواهد بود. یا اگر می خواهید در آینده در یکی از این نقش‌ها در Fabric کار کنید همچنین پایه ای عالی برای شما خواهد بود.

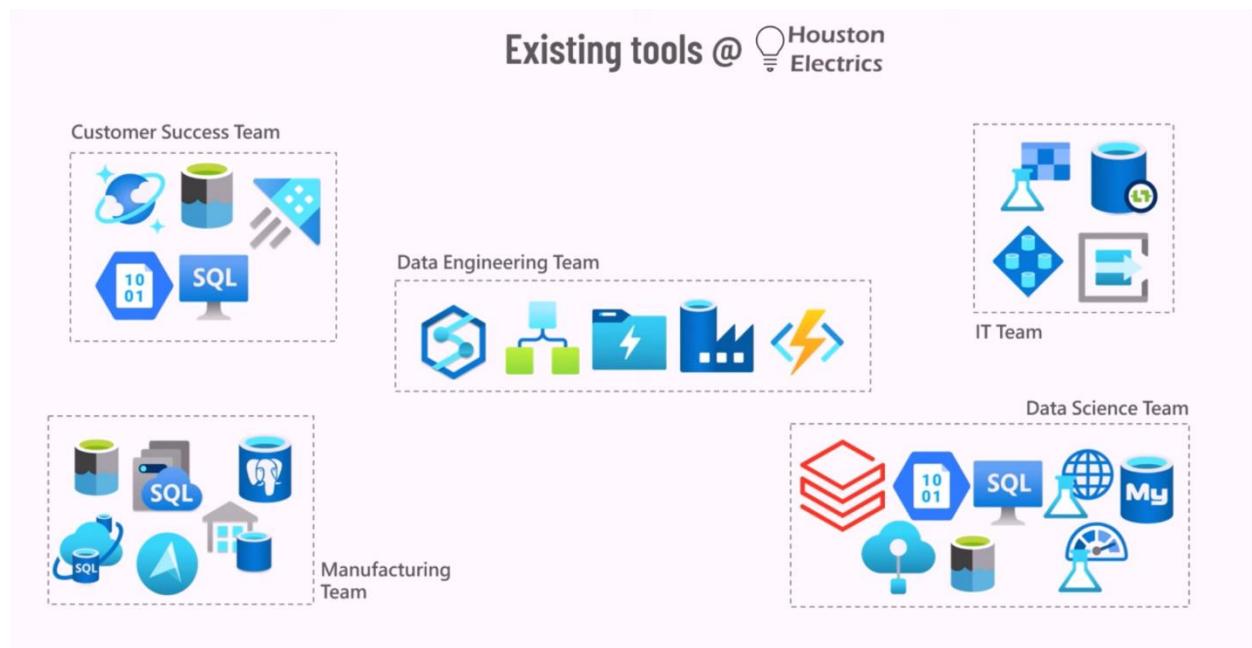
هیچ دانش پیش فرضی از فابریک نیازی ندارید.

خب برای درک اینکه چرا Fabric چنین فناوری تحول آفرینی است، به درک درست مشکلاتی که فابریک برای حل آنها ساخته شده است کمک می کند و برای نشان دادن این موضوع میخواهیم داستان هیوستون الکتریک را برای شما تعریف کنیم.

اکنون هیوستون الکتریک یک خرده فروش آنلاین کالاهای الکتریکی خیالی است که در ایالات متحده مستقر است اما قطعات الکتریکی را به سرتاسر جهان ارسال می کند. این شرکت به بیش از 400 کارمند افزایش یافته است تا به تقاضای رو به رشد خدمات ارائه دهد و شرکت به نوآور بودن خود افتخار می کند و از داده ها برای تصمیم گیری بهتر استفاده می

کند. در سال 2015، هیوستون الکتریک اولین مدیر ارشد داده خود را استخدام کرد، که یک برنامه تحول دیجیتال را برای استفاده بهتر از فناوری برای هدایت رشد کسب و کار راه اندازی کرد.

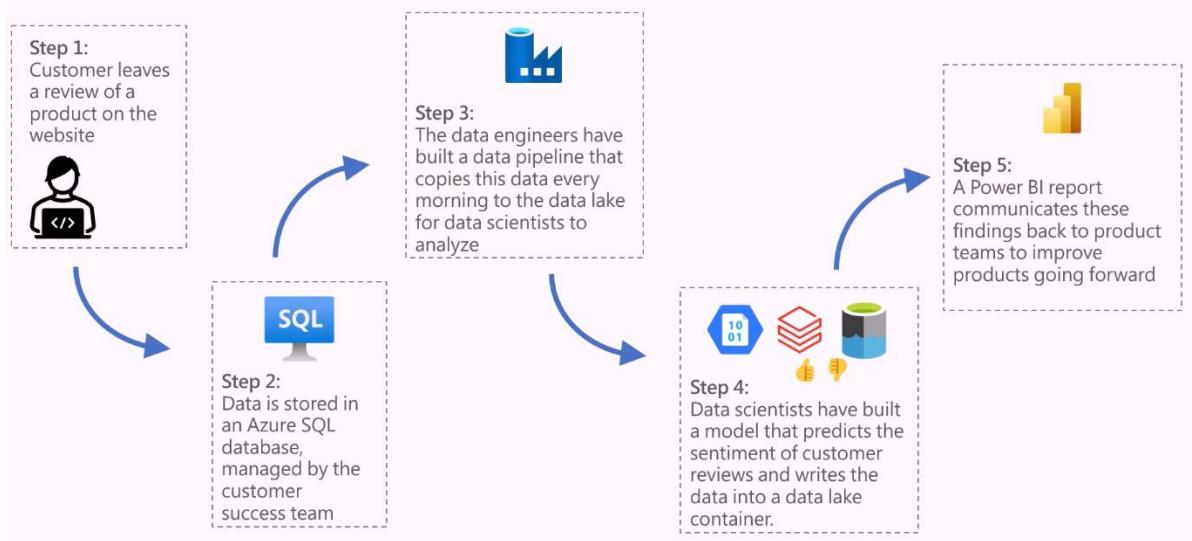
محور اصلی این استراتژی، ایجاد یک بخش داده مرکزی با بیش از 50 کارمند بود و این شرکت روی تعدادی از فناوری های ابری در خدمات وب Azure، Amazon سرمایه گذاری کرده است و آنها شروع به استفاده از Power BI کرده اند. با نگاهی به آنچه شرکت امروزه استفاده می کند، می بینیم که این شرکت حداقل از نظر فنی به بخش های اصلی زیر تقسیم شده است. هر بخش دارای تعدادی فناوری داده است که از آنها برای ذخیره، مدیریت و تجزیه و تحلیل داده های خود استفاده می کند. و اگر در اینجا با همه لوگوهای روی صفحه آشنا نیستید، نگران نباشید.



نکته اینجاست که چشم انداز داده ها در شرکت طی هشت سال گذشته به طور ارگانیک کمی رشد کرده است و چشم انداز استراتژیک کمی دارد. این یک وضعیت رایج است که بسیاری از شرکت ها در آن قرار می گیرند. هر بخش از ابزارهایی که ترجیح می دهد استفاده میکند و هر بخش برای انجام کار خود به داده های بخش های دیگر متکی است. به عنوان مثال، تیم موفقیت مشتری پایگاه داده بررسی سفارشات را مدیریت می کند، که تمام نظراتی را که مشتریان درباره محصولات خود می گذارند، پیگیری می کند.

اکنون بباید به یک workflow خاص نگاه کنیم که از این پایگاه داده بررسی سفارشات استفاده می کند تا ببینیم چگونه کار می کند و نحوه عملکرد شرکت را ببینیم.

Existing customer reviews workflow



وقتی مشتری محصولی را خریداری می کند، یک بررسی در وبسایت می گذارد و داده های بررسی در یک Azure SQL ذخیره می شود. پایگاه داده ای که توسط تیم موافقیت مشتری مدیریت می شود و تیم مهندسی داده و شرکت یک خط لوله ایجاد کرده اند که هر روز این داده ها را در **Data lake** کپی می کنند تا **Data science** ها تجزیه و تحلیل کنند. سپس **Data science** ها این داده های بازبینی را انجام می دهند و تجزیه و تحلیل احساسات را روی آن انجام می دهند. اساسا بررسی کنید تا متوجه شوید مشتریان بر اساس متمنی که می نویسند، از محصولات خاصی خوششان می آید یا از آنها متنفر هستند و سپس این احساسات در یک محفظه داده های **Lake** نوشته می شوند و در نهایت تیم **BI** یک گزارش **Power BI** ایجاد کرده است که این یافته ها را برای کمک به تیم های محصول ارسال می کند. آنها محصولات را در آینده بهبود می بخشنند. اکنون آنها این **Workflow** را به نقطه ای رسانده اند که کار می کند اما راه اندازی آن در دسیرساز بود و نگهداری آن بسیار در دسیرساز است و داده ها فقط برای اینکه یک گردش کار در چهار مکان مختلف در چهار مکان مختلف پراکنده شده است.

فرمت های داده ای که برخی از آنها اختصاصی هستند، به این معنی که شما واقعاً نمی توانید به راحتی از ابزارهای دیگر به آن دسترسی پیدا کنید و این تنها یک گردش کاری است که شرکت بسیاری از خطوط لوله پردازش لوله را اجرا می کند، مانند کپی کردن داده ها بین بخش های مختلف و بین مکان های ذخیره سازی و محصولات مختلف.

Existing tools @ Houston Electrics



Customer Success Team



100010111

Data Engineering Team



100010111

100010111

100010111

Manufacturing Team



Data Science Team



100010111

Existing tools @ Houston Electrics



Customer Success Team



100010111
Data Engineering Team



100010111
100010111

IT Team

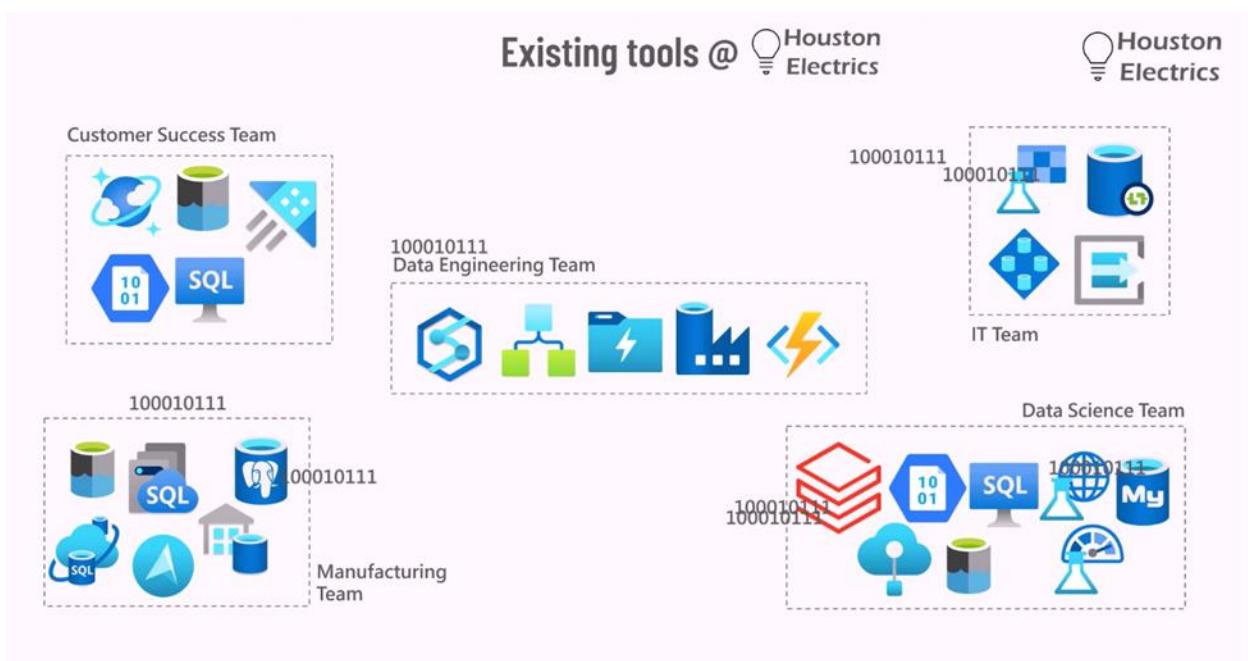
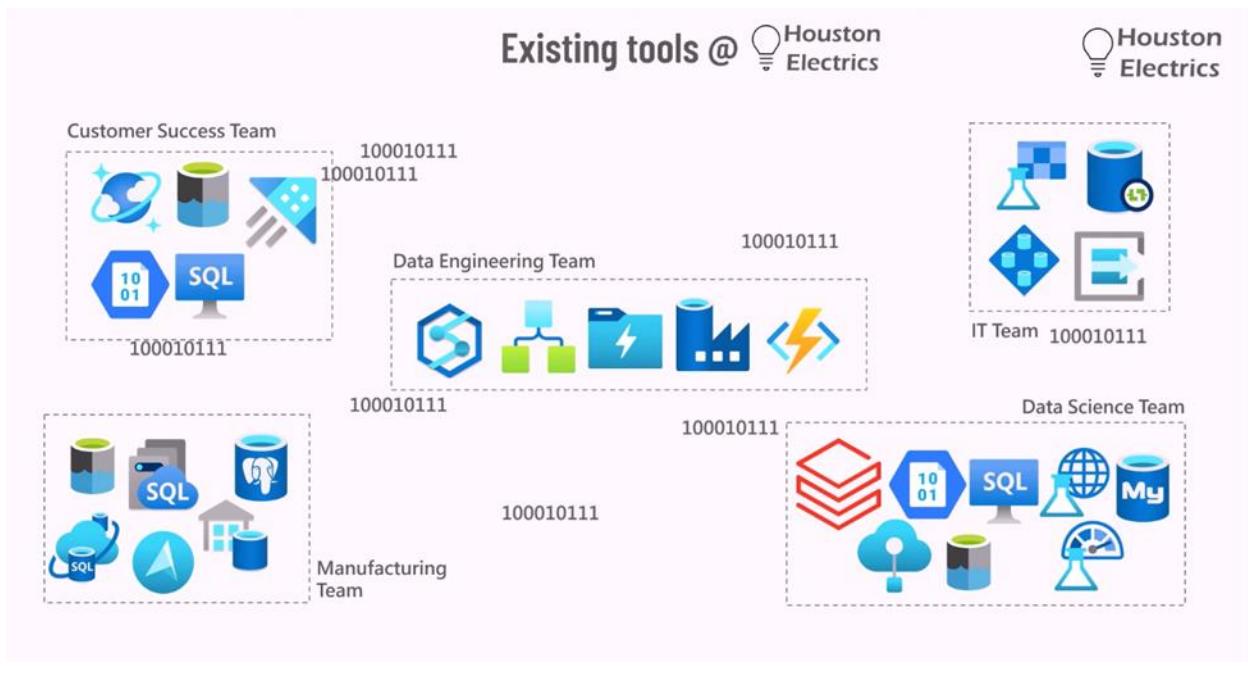


100010111

Manufacturing Team

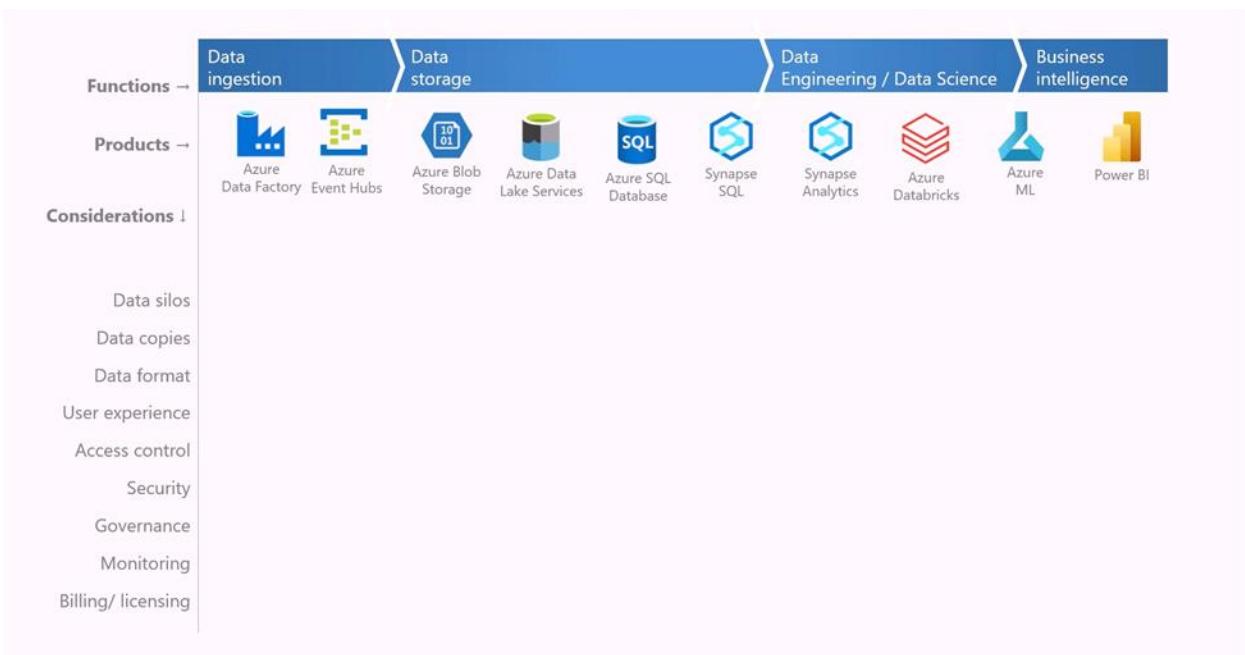


100010111



چشم انداز داده ها آنقدر رشد کرده است که تعمیر و نگهداری این سیستم ها و فرآیندهای موجود تقریبا به یک کار تمام وقت تبدیل می شود، بنابراین برای تجزیه و تحلیل آنچا ممکن است در اینجا اشتباہ شده باشد، اجازه دهید برخی از ایزارهایی را که شرکت در حال حاضر تحت این چهار سطح گسترده جذب داده، ذخیره سازی داده، مهندسی داده و علم داده و هوش تجاری استفاده می کند، ترسیم کنیم.

اکنون بباید وضعیت فعلی را با توجه به این ۹ ملاحظه که در سمت چپ فهرست شده است، تجزیه و تحلیل کنیم.



برای تجزیه و تحلیل آنها با تعدادی از کارمندان کلیدی شرکت میتوانیم صحبت کنیم. بنابراین اگر شما با مدیر ارشد داده صحبت می کنید او به شما خواهد گفت که معماری داده ها به طور ارگانیک رشد کرده است و اکنون سیلوهای داده در همه جا وجود دارد، من مدیر ارشد داده هستم و نمی خواهم مدیر ارشد یکپارچه سازی باشه. بنابراین به جای تمرکز بر زمان خود روی اینکه چگونه شرکت می تواند ارزش بیشتری از داده ها ایجاد کند، در اینجا نقش روی یکپارچه سازی سیستم متمرکز است که برای او، تیمش و برای شرکت دردرس بزرگی است.



IF YOU SPEAK TO...

Chief Data Officer

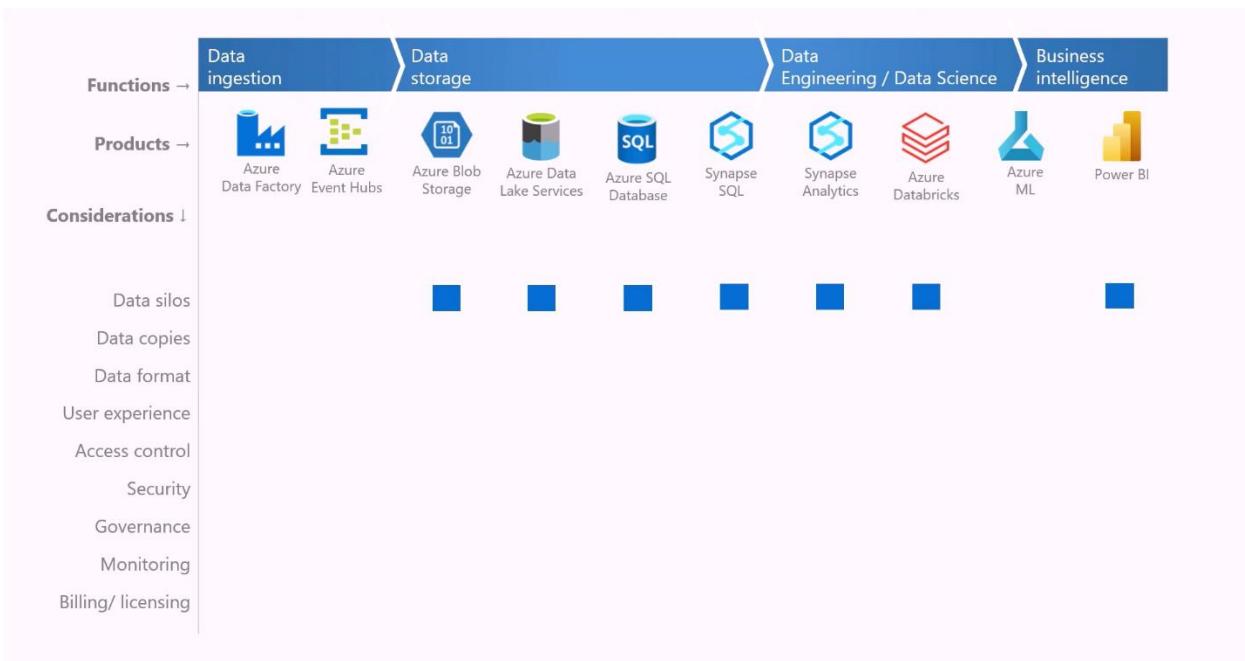
"

The architecture has grown organically and now there are data silos all over the place.

I am the Chief Data Officer and don't want to be the Chief Integration Officer

"

بنابراین در نقشه بصری خود در اینجا میتوانیم بلوک های مختلف را ترسیم کنیم که همه ی سیستم های مختلف را نشان می دهد که سیلوهای داده متفاوتی را ایجاد می کنند.



سپس اگر با مدیر مهندسی داده صحبت کنیم، او به ما می‌گوید که ما صدھا خط لوله را حفظ می‌کنیم که داده‌ها از بین تعداد زیادی از ذخیره‌گاه‌های داده مختلف برای بخش‌های مختلف کپی می‌کنیم و این کثیف و برای دور زدن است. تمام این سیلوهای داده در داخل و بین بخش‌ها و محصولات داده، تیم مهندسی داده‌ها خطوط لوله داده‌ای را ایجاد کرده‌اند که داده‌ها را از اینجا به آنجا کپی می‌کنند و اکنون آنها به عمق خط لوله هستند که به طور منظم از کار می‌افتد، زیرا مدیریت بسیار زیادی وجود دارد و در یک تیم کوچک نگهداری کنید.

IF YOU SPEAK TO...

Data Engineering Lead:

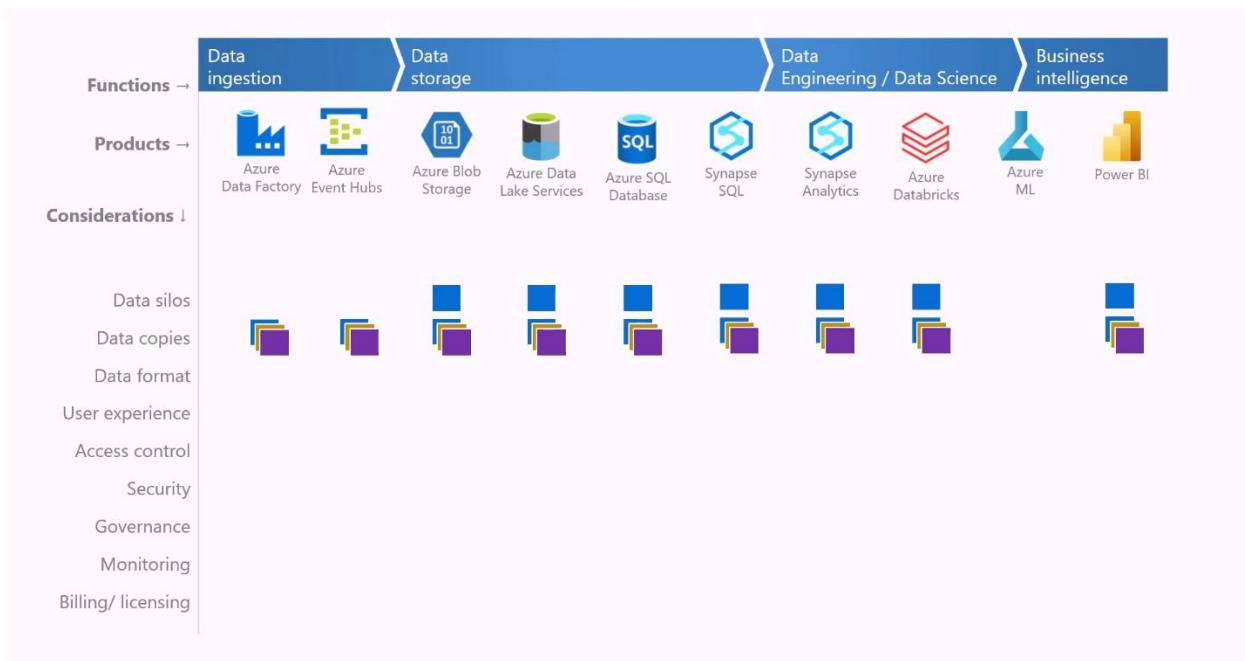
"

*We're maintaining hundreds of pipelines
copying data between lots of different
data stores for different departments.*

It's messy

"

بنابراین در نمودار ما اکنون می‌توانیم همه نسخه‌های مختلف بین هر سیستم و محصول را ببینیم و چگونه این باعث سردرگمی و نگهداری و مشکلاتی در کیفیت داده‌ها می‌شود، همچنین فقط تیم مهندسی داده نیست که باعث ایجاد مشکلاتی می‌شود.



تصور کنید که یک توسعه دهنده Power BI هستید، چگونه می دانید به کدام مجموعه داده ها باید به یک سازمان متصل شوید و چگونه می توانید بدانید که آیا آنها به روز هستند و آیا می توانید به آنها اعتماد کنید؟

تیم Data science با مشکلات مشابهی روبرو می شود که آنها به آن اشاره می کنند که ما داریم. داده ها در قالب های بسیار در مکان های مختلف پراکنده شده اند و روزها طول می کشد تا مجموعه داده تمیز را فقط برای شروع تجزیه و تحلیل بدست بیاورم و حتی در آن زمان نمی دانم می توانم به داده ها اعتماد کنم یا خیر.

IF YOU SPEAK TO...

Data Scientist:

"

*We have data scattered in so many formats in different places,
it takes me days just to get clean datasets just to begin an
analysis. Even then, I don't know if I can trust the data.*

"

علاوه بر آن، فقط این واقعیت که سیستم های بسیار زیادی وجود دارند و همه آنها متفاوت هستند یک سربار شناختی برای این تیم ایجاد میکند. آنها اضافه کردند که من باید پیچیدگی های بسیاری از فناوری های داده مختلف را باید می گرفتم و هر کدام متفاوت است، ما یک استارتر جدید داشتیم و سه ماه طول کشید تا آنها را در تمام پلتفرم های داده متفاوتی که استفاده می کنیم ارتقا دهیم.

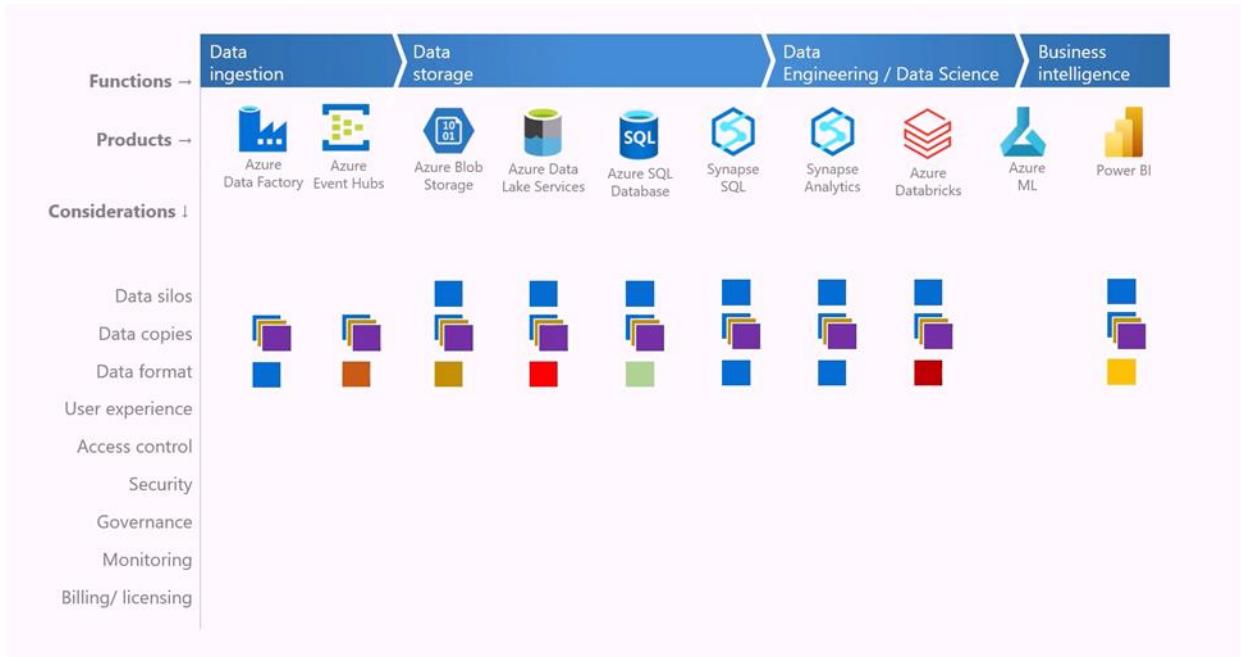
IF YOU SPEAK TO...

 Data Scientist:

We have data scattered in so many formats in different places, it takes me days just to get clean datasets just to begin an analysis. Even then, I don't know if I can trust the data.

I had to learn the intricacies of many different data technologies, and each one is different. We had a new starter. It took 3 months to upskill them in all the different platforms we use.

بنابراین بباید این مشکلات را به تصویر خود اضافه کنیم.



اشاره Data scientists کردند که هیچ یکنواختی در قالب های داده در سرتاسر سازمان وجود ندارد، که باعث می شود همه چیز به یک قالب مشترک تبدیل شود. برای شروع تجزیه و تحلیل مورد نیاز است. علاوه بر این، آنها باید هر روز روی پنج، شش، شاید هفت محصول داده کار کنند، که هر کدام منحنی یادگیری، رابط کاربری، تجربه کاربری خاص خود

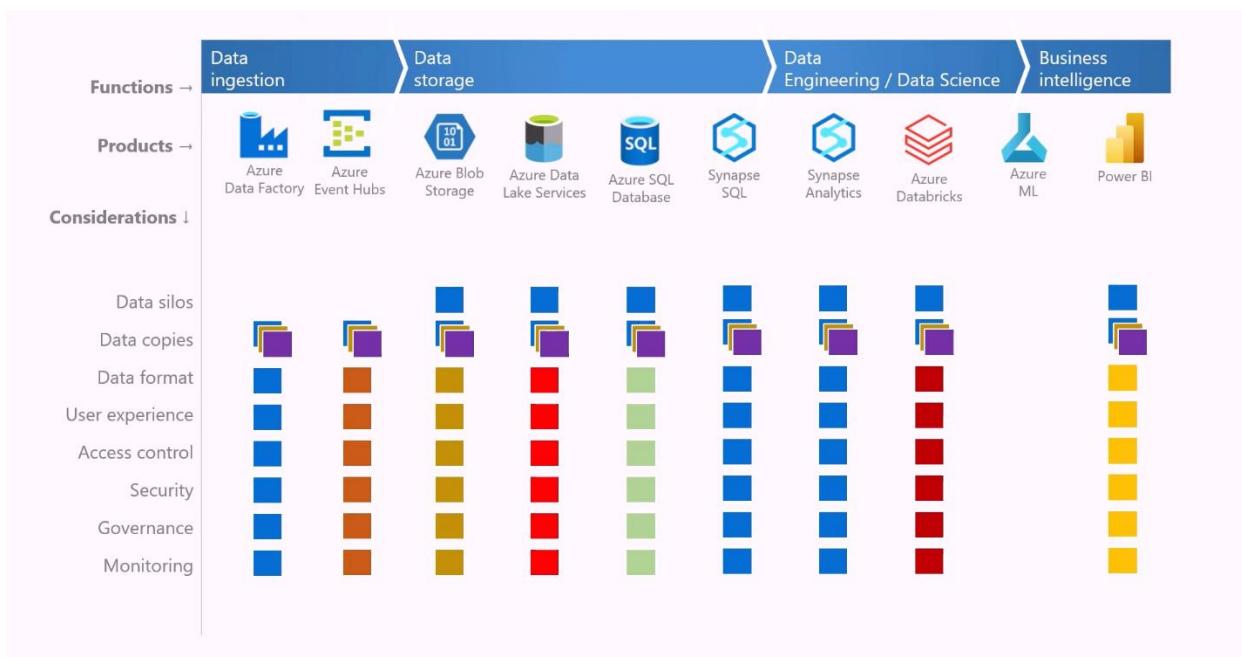
را دارند. و این ناکارآمد است، به خصوص برای وصال های جدید. بنابراین بلوک های رنگی مختلف در اینجا نشان می دهد که حکم‌گوئه تحریه کاربر در هر یک از این احتمالات، که شرکت استفاده می کند متفاوت است.

بعدی، مدیر IT نیز خوشحال نیست. او می‌گوید ما از سیستم‌های بسیار زیادی استفاده می‌کنیم. همه دارای پروفایل‌های امنیتی و الزامات متفاوتی برای حفظ امنیت داده‌ها و انتقال داده‌ها در حالت امن هستند و این یک کالوس است.



"We're using too many systems, all have different security profiles and requirements to keep data at-rest and in-transit secure. It's a nightmare"

این درست است که فقط افراد حاضر در این سیستم ها نیستند.



ابزارهایی که برای مدیر فناوری اطلاعات و تیمش که مسئولیت مدیریت دسترسی به هر یک از این ابزارها را به طور مستقل این می‌کند و داده‌های ذخیره شده در آنها را کنترل می‌کند و سپس نظارت و نگهداری سیستم‌ها را به طور کلی و به گونه‌ای که اینطور نیست، دشوار است. به اندازه کافی بد است که او همچنین هر ماه با مدیر مالی تماس می‌گیرد تا در مورد لایحه‌ی Azure شکایت کند. او افزود، من می‌ترسم که صورتحساب Azure خود را هر ماه دریافت کنم. این بسیار غیر قابل پیش‌بینی و گاهی ترسناک است که هر محصول داده ساختار قیمتی خاص خود را دارد، بنابراین پیش‌بینی چگونگی آن دشوار است. ماه به ماه هزینه زیادی از ما دریافت می‌شود.

IF YOU SPEAK TO...



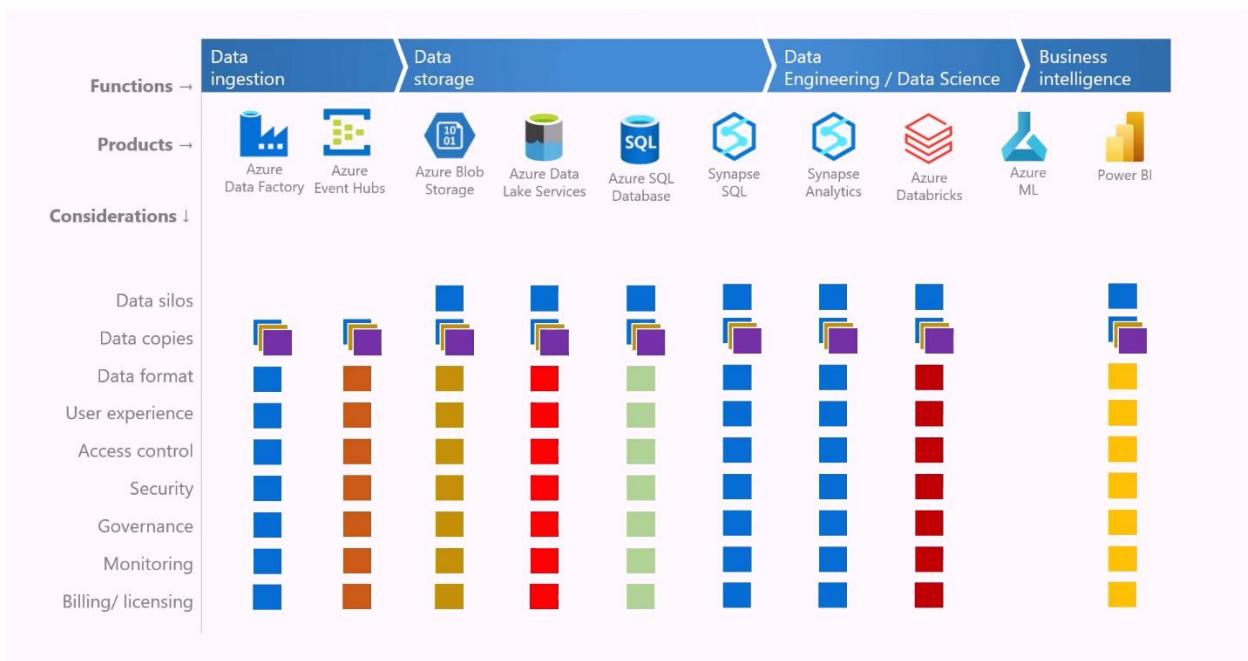
IT Director

"

I dread getting our Azure bill every month. So unpredictable, and sometimes scary. Each data product has their own pricing structure so it's difficult to predict how much we will be charged month-to-month.

"

بنابراین اکنون صورتحساب و مجوز را به تصویر خود اضافه می‌کنیم و می‌بینیم که هر محصول مجوز متفاوتی دارد و ساختار صورتحساب متفاوتی دارد که مدیریت شرکت‌ها را واقعاً دشوار می‌کند.



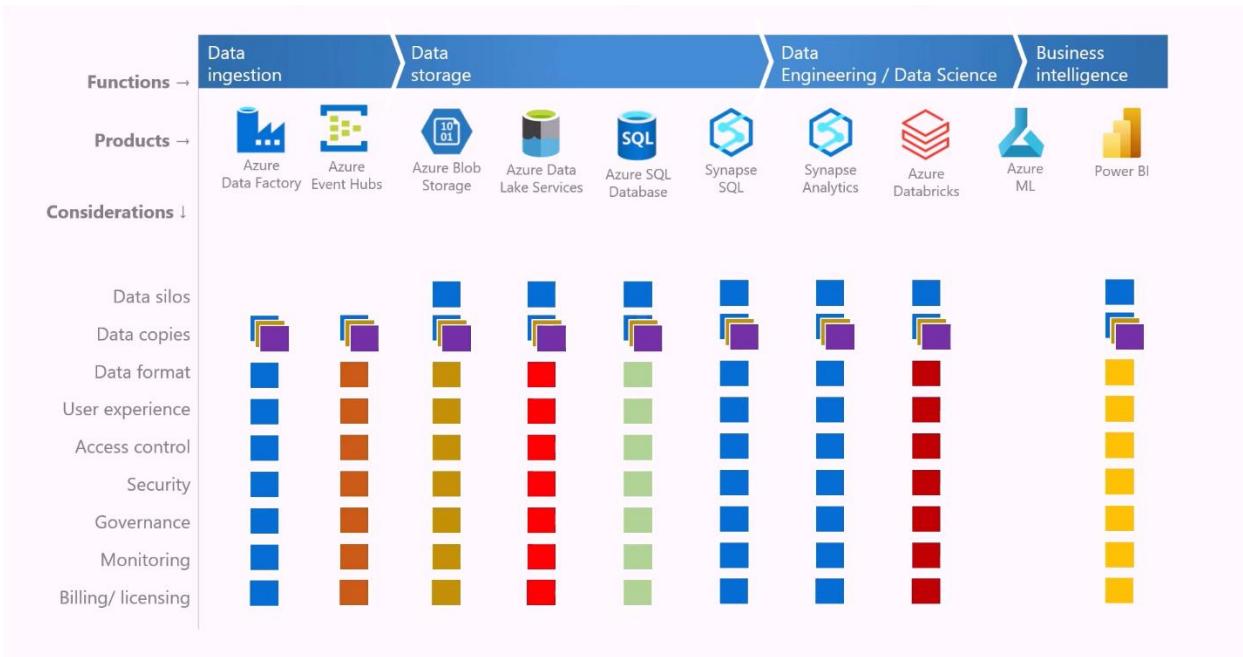
در نهایت، این کاربر بیجارد داشبورد است. آنها واقعاً چیز زیادی در مورد همه پلاوهایی که در پشت صحنه اینجا می‌گذرد نمی‌دانند. اما آنها می‌توانند احساس کنند که چیزی درست نیست. آنها می‌گویند، من مطمئن نیستم که بتوانم به داده‌هایی که در اینجا به من ارائه می‌شود اعتماد کنم. همیشه واقعیت را منعکس نمی‌کند و حق دارند عصبی باشند.

IF YOU SPEAK TO...

Dashboard user:

*I'm not sure if I can trust the data
that is being presented to me –
doesn't always reflect reality*

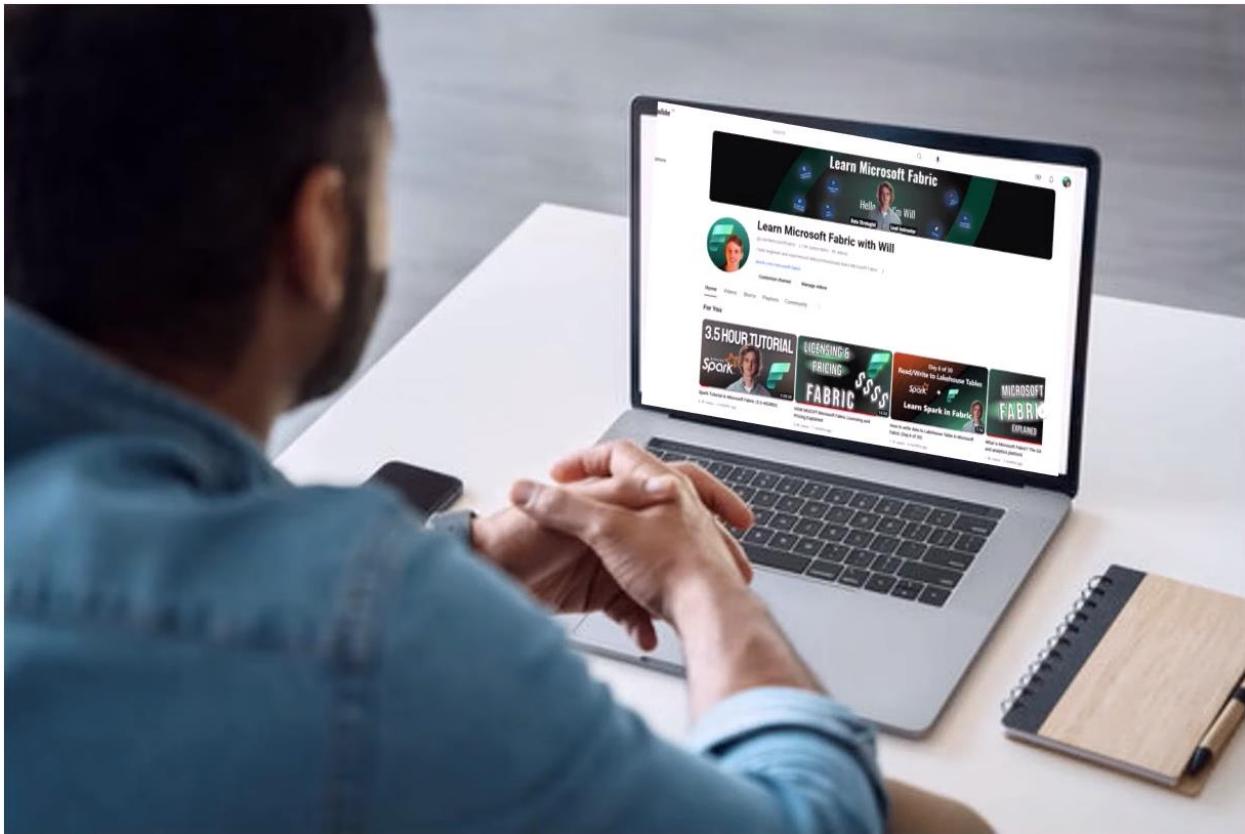
تمام این پیچیدگی به این معنی است که تیم‌های داده نتوانسته اند واقعاً کنترل داده‌ها و کیفیت داده‌ها را کنترل کنند و کاربران داشبورد در پایان فرآیند اعتمادشان را از دست می‌دهند و وقتی اعتمادشان به تجسم‌ها و داده‌هایی که به آنها ارائه می‌شود از دست می‌دهند، پس به هر حال همه اینها چه فایده‌ای دارند؟



پس بباید یک ثانیه در آنجا مکث کنیم

And pause...

و... در این لحظه بود که رئیس هیوستون الکتریک، مرد باهوشی که هست، داشت کanal یوتیوب Microsoft Fabric را تماشا می کرد. وای چه تصادفی! او درباره فابریک مایکروسافت یاد گرفت و می دانست که مایکروسافت Learn فابریک می تواند به حل مشکلاتی که آن ها به عنوان یک کسب و کار تجربه می کردند، کمک کند و حق داشت.



بیینید، در مایکروسافت فابریک، ما خانواده ای از محصولات داده ای را در دسترس داریم که به راحتی زیر چهار سطل قرار می گیرند که قبل آنها را بررسی کردیم.

1. Data ingestion (جذب داده ها)
2. Data storage (ذخیره سازی داده ها)
3. Data Engineering/Data Science (علم مهندسی داده ها)
4. Business intelligence (هوش تجاری)



اکنون این ابزارها، هنگامی که به صورت مجزا به آنها نگاه می شود، عملکردی مشابه ابزارهایی را که شرکت قبل استفاده می کرد، انجام می دهند.

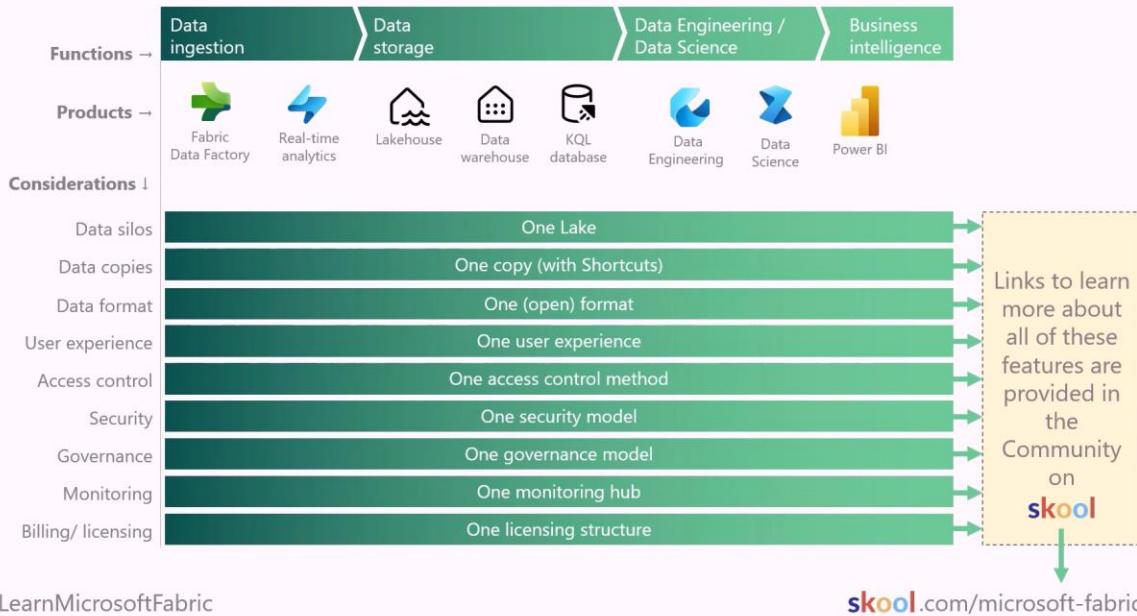


این یک امر رایج است که گاهی اوقات می شنوید که مردم در مورد فابریک مایکروسافت می گویند که او، این فقط یک تمرین بازاریابی مجدد است. آنها فقط یک نشان جدید بر روی فناوری موجود گذاشته اند. اما این دیدگاه اساساً به اشتباہ درک می کند که Fabric کاملاً از ابتدا ساخته شده است تا تمام مشکلاتی را که در اسلامیدهای قبلی دیدیم برطرف کند. این فقط یک تمرین بازاریابی نیست، بلکه یک بازنگری کامل و معماری مجدد در مورد نحوه مدیریت داده ها در سازمان شما است. و بباید منظورمان از آن را با جزئیات بیشتر بررسی کنیم. بنابراین در Fabric، تمام داده های شرکت شما در یک مکان قرار دارند و آن را OneLake می نامند و این مفهوم اساسی است که ما کمی بعد با جزئیات بیشتری در مورد اینکه چگونه OneLake سیلوهای داده را ریشه کن می کند.



1. داده های همه این محصولات مختلف در واقع در یک مکان ذخیره می شوند. و به همین دلیل، نیاز به ایجاد چندین نسخه از یک مجموعه داده را نیز از بین می برد.
2. یک اصل اساسی در Fabric این است که یک مجموعه داده فقط باید در یک مکان وجود داشته باشد و سپس با استفاده از یک ویژگی هوشمندانه به نام میانبر در سراسر Fabric ارجاع داده شود. بنابراین تمام داده های سازمان شما در OneLake ذخیره شود.
3. در OneLake، همه داده ها در قالبی یکسان به نام دلتا پارکت ذخیره می شوند. و این فرمت استانداردهای باز است. و این بسیاری از مشکلاتی را که قبل بر جسته کردیم حل می کند. مهمتر از همه، مشکل یکپارچه سازی داده ها را حل می کند تا Data scientists و Data engineers ممکن است همگی از ابزار های متفاوتی استفاده کنند، شاید از زبان های مختلفی که با آن ها آشنا هستند، اما در زیر سرپوش، همه آنها روی داده های مشابهی کار می کنند که در قالب یکسان است. و آنها ساعت ها و روز ها را برای به دست آوردن تبدیل این داده ها به قالب مناسب تلف نمی کنند تا بتوانند تجزیه و تحلیل خود را آغاز کنند.
4. در مرحله بعد، Fabric یک تجربه کاربری یکپارچه ارائه می دهد. قبل شنیده بودیم که مردم از ورود به چندین پلتفرم مختلف که هر کدام ظاهر و احساس متفاوتی دارند خسته شده اند. هنگامی که در Fabric از طریق یک پورتال وب وارد سیستم می شوید و تجربه کاربری به گونه ای طراحی شده است که شبیه تجربه مايكروسافت 365 باشد. بنابراین یک بار وارد سیستم می شوید و می توانید به هر تجربه Fabric بروید و هر تجربه ظاهر و احساسی مشابه دارد. این بدان معنی است که می توانید زمان بیشتری را صرف فکر کردن در مورد اینکه چگونه می خواهید از داده های خود ارزش کسب کنید، صرف کنید. و نگران نحو پیمایش در برنامه نباشید. و برای افرادی که Fabric را مدیریت می کنند، مزایای زیادی نیز وجود دارد.
5. برای شروع، کنترل دسترسی و امنیت در Fabric به شدت ساده شده است. یک روش کنترل دسترسی و یک مدل امنیتی در تمام ابزارها و تجربیات اعمال شده است. و دسترسی به این منابع اصولاً از طریق فضاهای کاری مدیریت می شود، که فقط مجموعه ای از آیتم های Fabric هستند که برای مرز امنیتی خاصی که می خواهید تنظیم کنید، منطقی است. و از آنجایی که همه داده های ما در یک Lake زندگی می کنند، مدیریت داده ها و قابلیت کشف آنها کار بسیار ساده تری می شود.
6. و در واقع، Fabric دارای ویژگی های بسیاری برای مدیریت مجموعه داده های شما است.
7. مورد بعدی واقعاً مهم است Fabric دارای یک هاب ناظارت داخلی است که تمام فعالیت های Fabric را در تجربیات شما نظارت می کند. بنابراین اگر می خواهید تمام خطوط لوله داده های مختلف، نوت بوک های اجرا شده و هر نوع فرآیندی را که در Fabric اجرا می کنید، نظارت کنید، فقط یک مکان برای جستجو دارید.
8. و در آخر، صدور صورت حساب و مجوز در Fabric یکپارچه است، به این معنی که وقتی مقداری ظرفیت Fabric را برای سازمان خود خریداری می کنید، بلاfacله به همه ویژگی ها و موارد دسترسی پیدا می کنید. دیگر مجوز یا ساختار صورتحساب جداگانه ای برای همه محصولات مختلف فهرست شده در بالا وجود ندارد. اکنون سطوح مختلفی از ظرفیت وجود دارد که می توانید بسته به شدت استفاده خود خریداری کنید. و چیزی که در اینجا باید در نظر داشت این است که اگر شرکت شما در حال حاضر برای ظرفیت پریمیوم BI Power پرداخت می کند، یک ظرفیت فابریک F64 دریافت می کنید که ظرفیت بسیار زیادی در فابریک به صورت رایگان با ظرفیت پریمیوم BI Power شما است. این بدان معناست که امروز می توانید بدون پرداخت هزینه اضافی، هر افلام فابریک ای را ایجاد و استفاده کنید. و همه این مواردی که در اینجا به آنها اشاره کردیم به نوعی موضوعی هستند که می توانیم به جزئیات بیشتری در مورد آنها بپردازیم.

HOW FABRIC IS DIFFERENT



اگر وبسایت گذاشته شده در عکس را گوگل کنید در آنجا می توانید با جزئیات بیشتر در مورد هر یک از این ویژگی های خاص در Fabric اطلاعات کسب کنید.

خوب، پس اکنون ما شروع به درک قدرت Fabric کرده ایم و شما در مورد One Lake و تجربیات و موتورهای محاسباتی و اینکه این چیزها وقتی با هم کار می کنند چقدر قدرتمند هستند شنیده اید. بباید به نحوه عملکرد آن با جزئیات بیشتر نگاه کنیم.

Ok so now we're starting to understand
the power of Fabric...

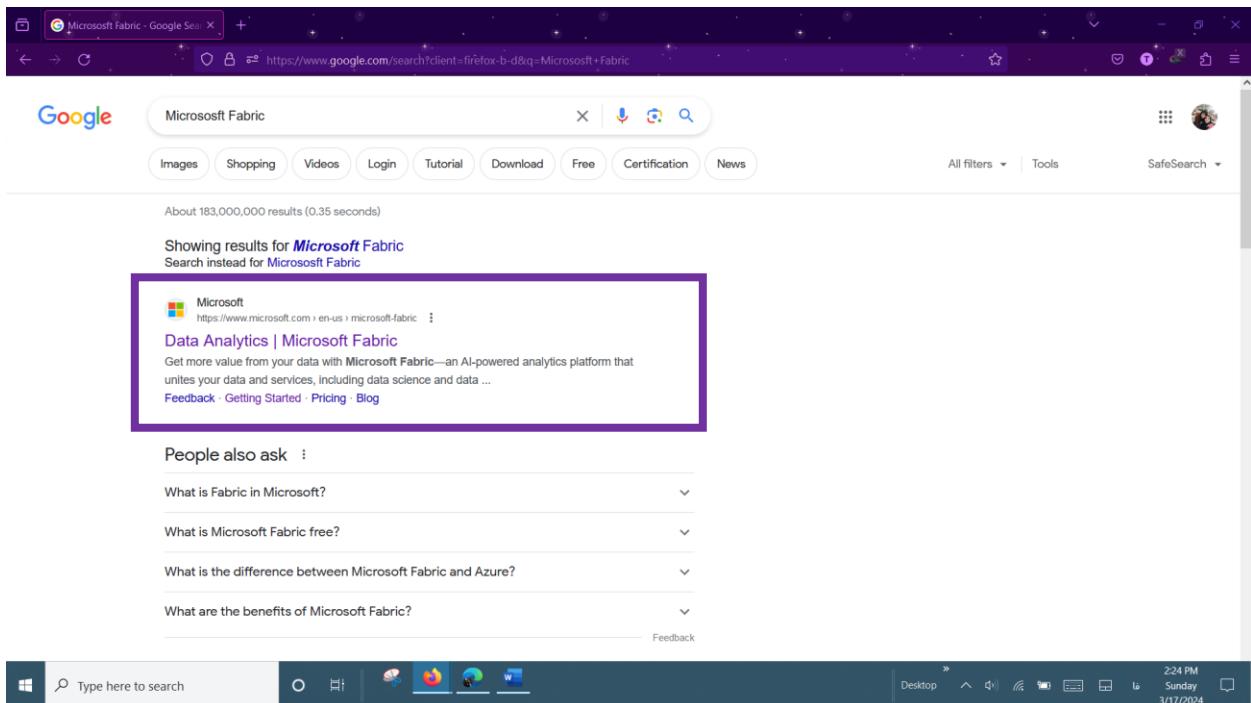
Well, let's explore in more details what is included
and how it works...

فابریک چیست و تمام عناصر مختلف فابریک مایکروسافت را بررسی میکنیم که چگونه با استفاده از آنها میتوانیم به نوعی راه حل منسجم برای سازمان خود یا برای خود جمع آوری کنیم.

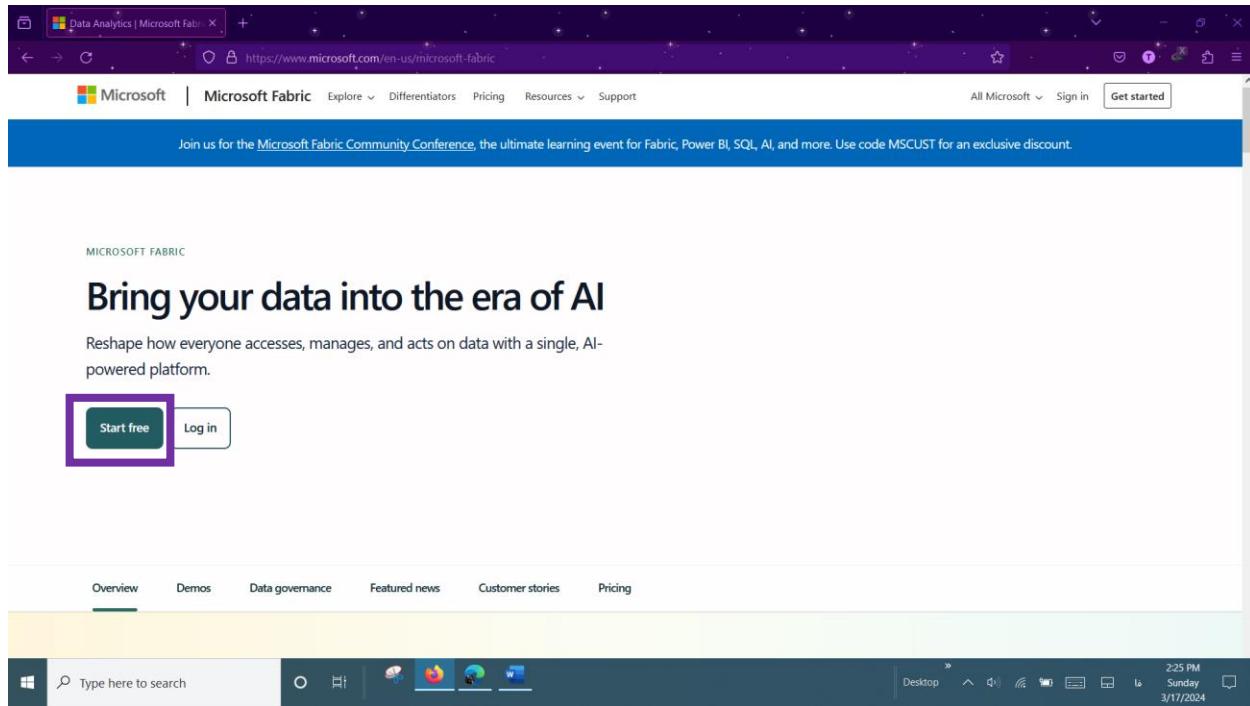
فابریک در سال 2023 منتشر شده است.

ورود به وبسایت Microsoft Fabric

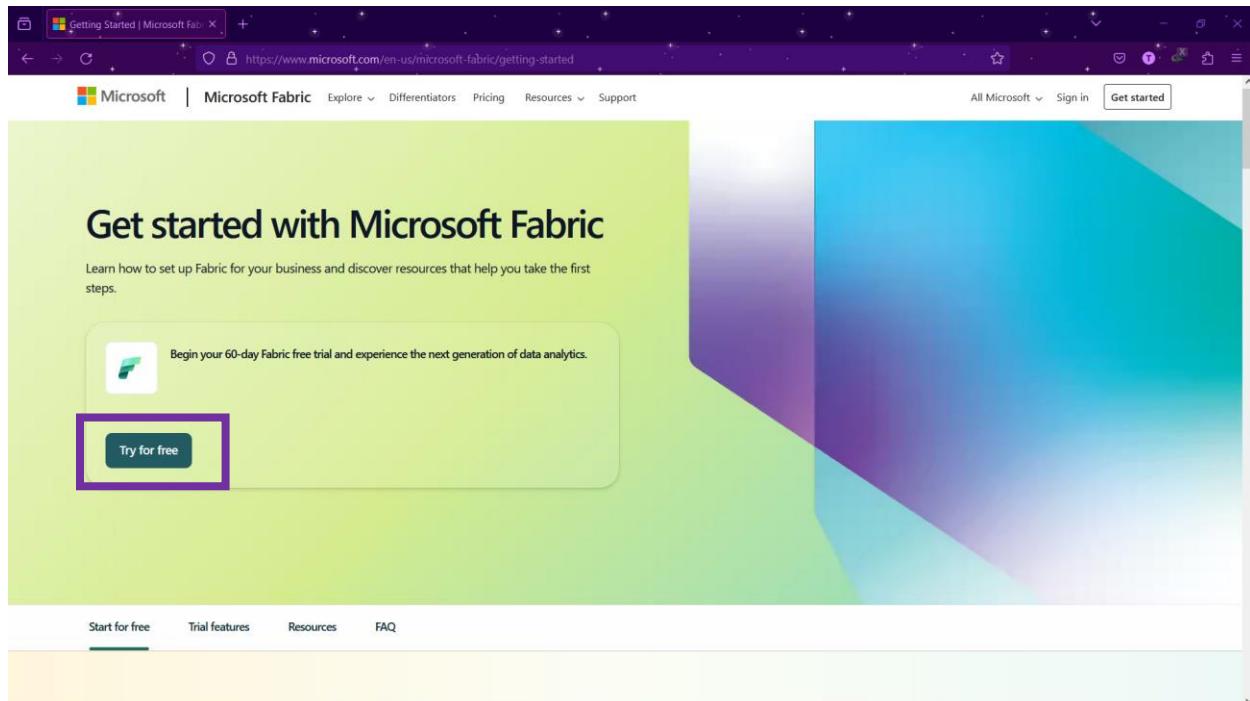
توی گوگل سرچ میکنیم Microsoft Fabric تا به نصب آن بپردازیم. وارد وبسایت خود مایکروسافت میشویم.



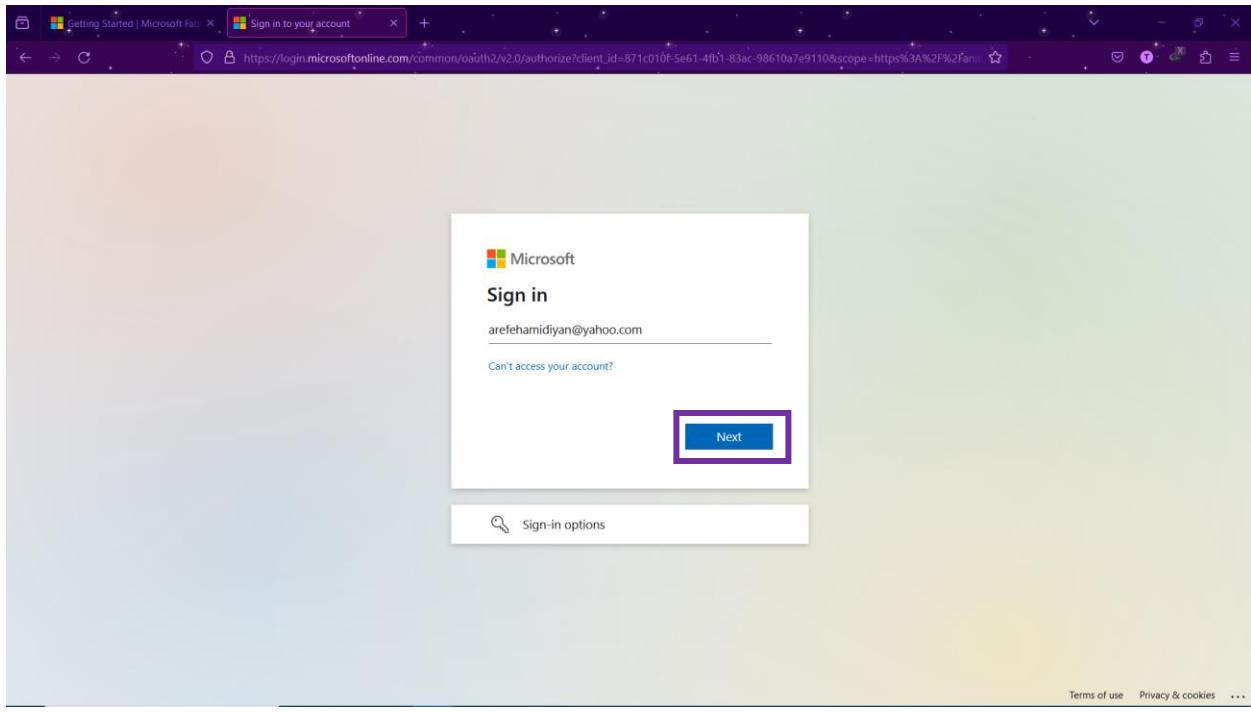
نرم افزار رو نمیتوانیم نصبش کنیم و باید از وبسایت کارمون رو انجام بدیم. روی start free کلیک میکنیم.



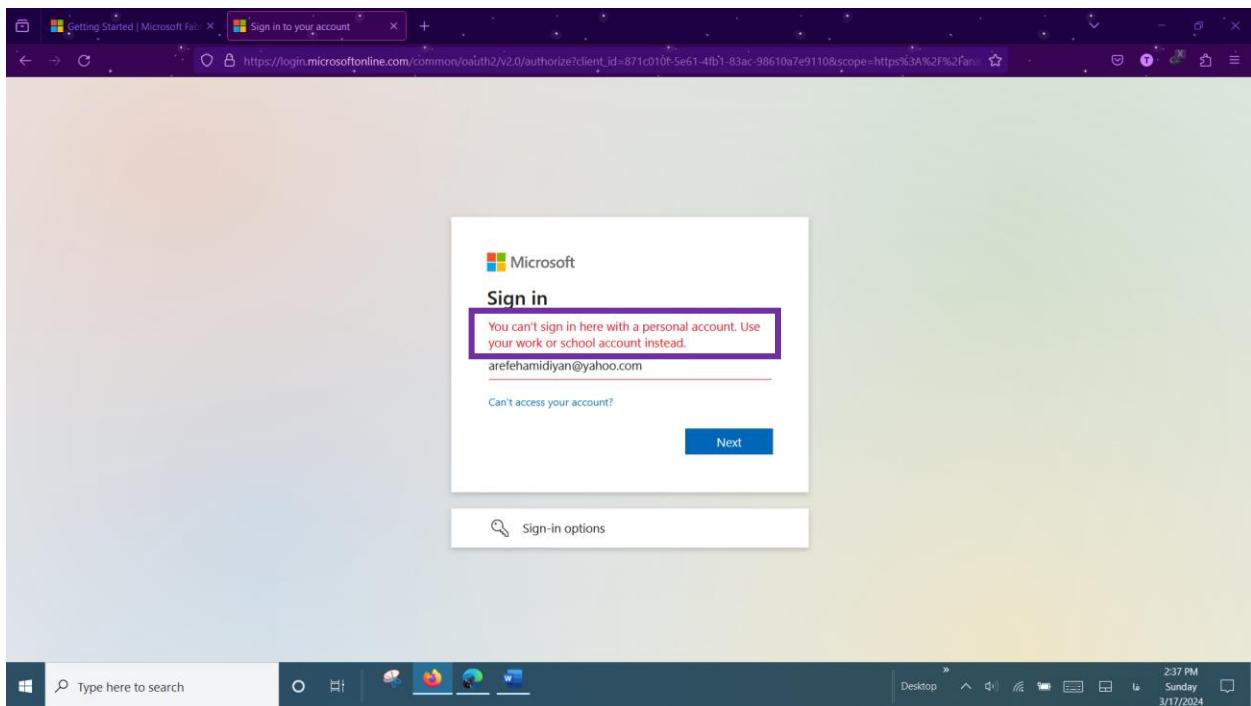
نرم افزار رو نمیتوانیم نصبش کنیم و باید از وبسایت کارمون رو انجام بدیم. روی start free کلیک میکنیم.
در اینجا هم Try for free رو انتخاب میکنیم.



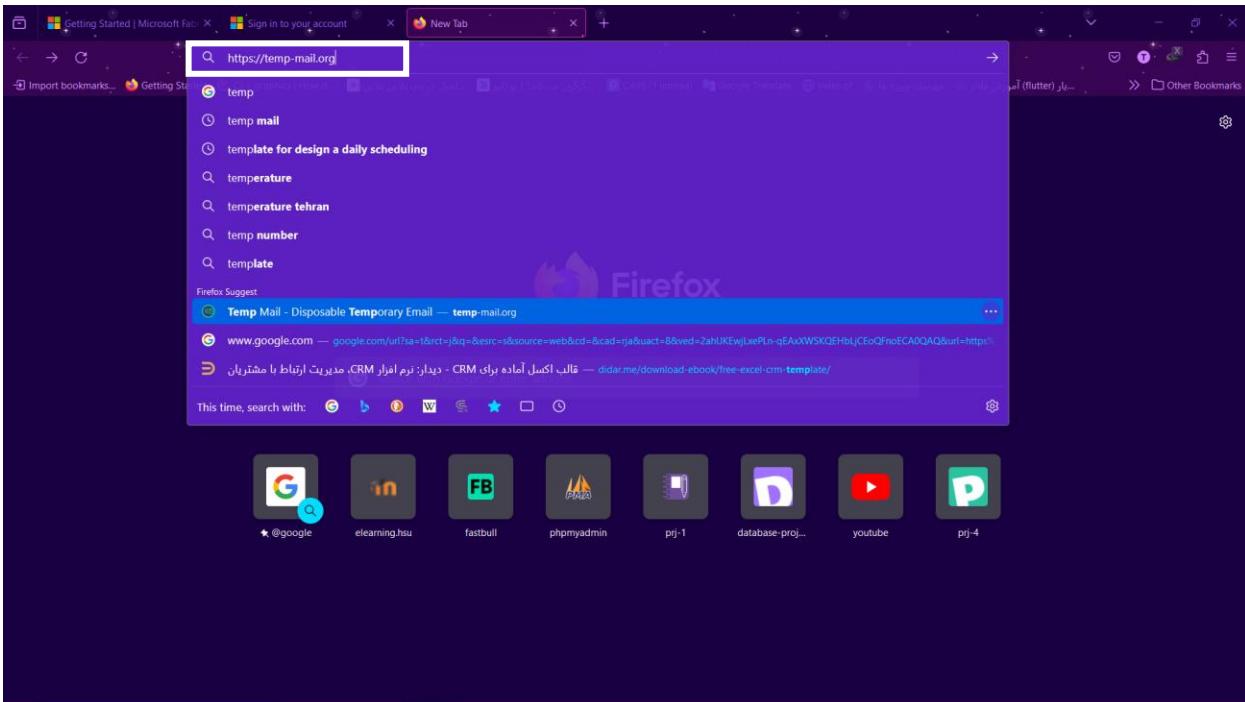
اینجا از مون میخواد که وارد اکانت مایکروسافتمن بشیم. ایمیلمون رو وارد می کنیم. Next رو میزنیم.



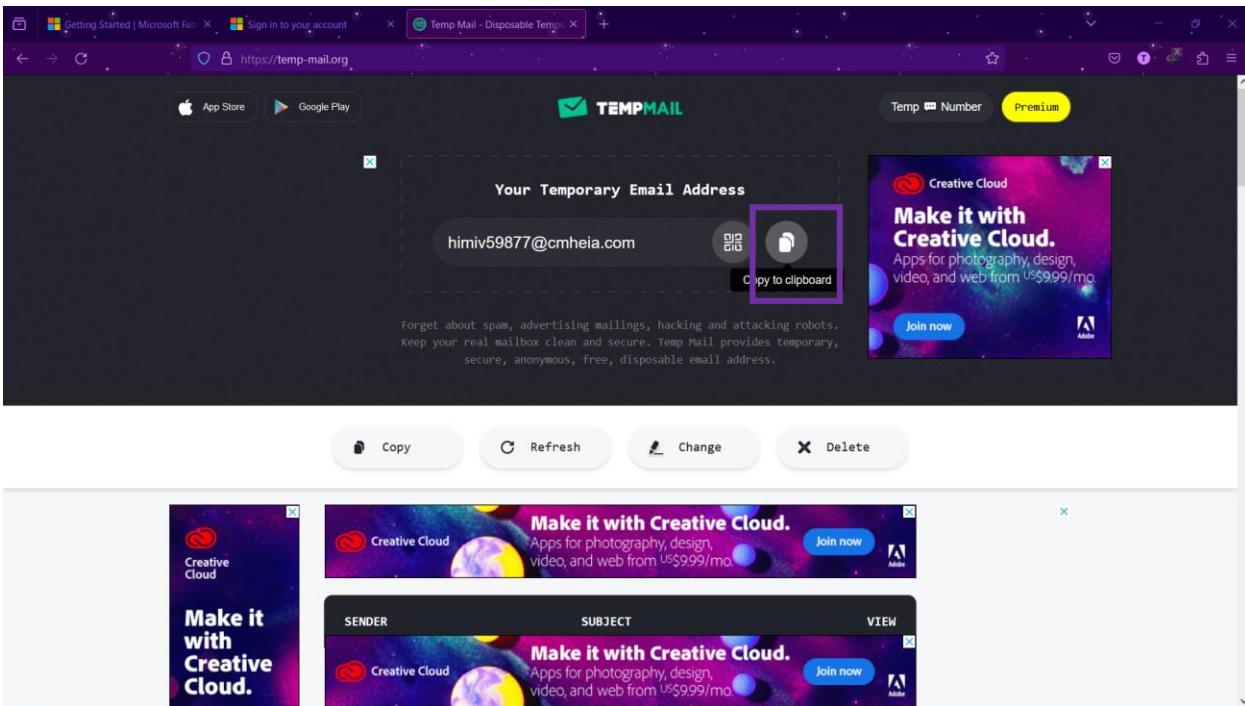
در اینجا اروری به ما نمایش میدهد که باید حتما برای ورود ایمیل کاری باشے برای اکانت مایکروسافت و ایمیل شخصی قابل قبول نیست.



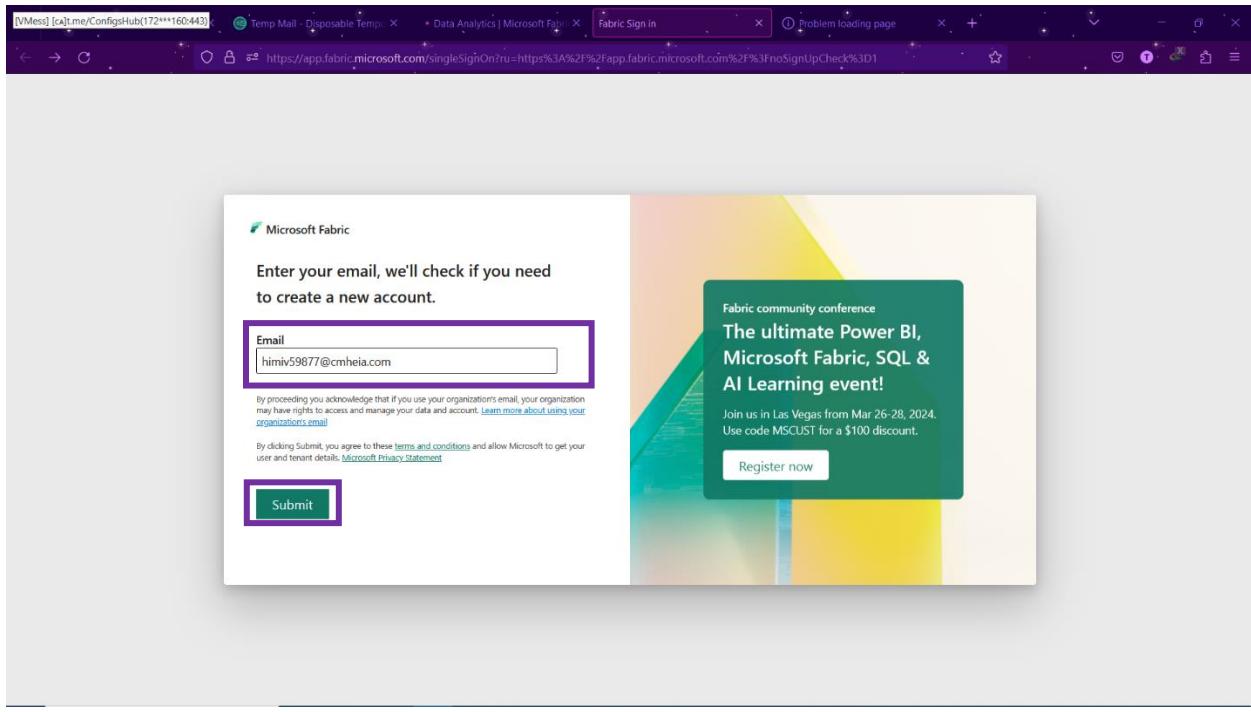
چون برای ما ایمیل کاری معتبر مقدور نبود ما از یک ایمیل فیک استفاده میکنیم. وارد سایت temp mail.org میشویم.



یک ایمیلی به ما میده.



ما اون رو وارد مایکروسافت میکنیم.



رو میزنیم.



به این صفحه که وارد شدیم از ما میپرسه که ایمیلمون شخصیه یا کاری. ما روی کاری میزنیم. چون شخصی قبول نمیکنه. Next رو میزنیم.

You've selected Microsoft Fabric free

1 Let's get you started

Looks like you need to create a new account.

Microsoft Fabric free is designed to be used by people collaborating within an organization, so your email will be visible to others who also use @cmheia.com email addresses to signup

For that reason, emails from shared email services like outlook.com should not be used.

What kind of email is himiv59877@cmheia.com?

I got it from my organization

It's my personal email

Next

By proceeding you acknowledge that if you use your organization's email, your organization may have rights to access and manage your data and account.
[Learn More](#)

2 Create your account

3 Confirmation details

خوب وارد صفحه ثبت نام حساب میشویم. اطلاعات شخصی از قبیل اسم، فامیل، شماره تلفن کاری، پسورد و رو وارد میکنیم.

2 Create your account

First name *

Last name *

Country or Region *

Business phone number *

Email

Enter a password to sign in to your account.

Create password *

Confirm password *

We've sent a verification code to himiv59877@cmheia.com. Enter the code to complete sign up.

Verification code * [Resend code](#)

I understand that Microsoft may contact me about my trial.

I will receive information, tips, and offers about Power BI, solutions for businesses and organizations, and other Microsoft products and services.

[Privacy Statement](#)

از temp mail 'inbox' کد رو برمیداریم.

The screenshot shows a web browser window with multiple tabs. The active tab displays a temporary email address: himiv59877@cmhei.com. The page includes a Creative Cloud advertisement and a message about the benefits of using a temporary email. Below the address, there are buttons for Copy, Refresh, Change, and Delete. A second tab shows an error page, and another tab is for Microsoft Fabric free sign-up.

برای شماره تلفن کاری هم یک شماره فیک از سایت زیر برمیداریم.

The screenshot shows a web browser window for Talkroute.com. The page features a search bar for 'Los Angeles, CA 310'. Below it, a grid of 12 virtual phone numbers is displayed, with the first one highlighted. The page includes a 'Start Free Trial' button, a 'Book Live Demo' button, and a 'New Client Special Offer' section. The URL in the address bar is https://talkroute.com/sign-up/us-virtual-number/

اطلاعات که کامل شد روی Next میز نیم.

The screenshot shows a Microsoft Edge browser window with the URL <https://signup.microsoft.com/get-started/signup?sku=a403ebcc-fae0-4ca2-8c8c-7a907fd6c235&email=himiv59877%40cmheia.com&ru=https://www.microsoft.com/fabric/>. The page is titled "Create your account" (Step 2). It contains fields for First name (Trinity), Last name (Ami), Country or Region (United States), Business phone number (13105553890), Email (himiv59877@cmheia.com), Create password, Confirm password, Verification code (490869), and Resend code. Below the verification code field is a note about receiving information from Microsoft. At the bottom, there is a checkbox for sharing information with partners and links to the Privacy Statement.

(2) Create your account

First name * Last name *

Country or Region *

Business phone number *

Email

Enter a password to sign in to your account.

Create password * Confirm password *

We've sent a verification code to himiv59877@cmheia.com. Enter the code to complete sign up.

Verification code * [Resend code](#)

I understand that Microsoft may contact me about my trial.
I will receive information, tips, and offers about Power BI, solutions for businesses and organizations, and other Microsoft products and services.
[Privacy Statement](#).

I would like Microsoft to share my information with select partners so I can receive relevant information about their products and services. To learn more, or to unsubscribe at any time, view the [Privacy Statement](#).

[VMess] [ca]t.me/ConfigHub(172***160:443) Temp Mail - Disposable Temp... Getting Started | Microsoft Fabric Microsoft Fabric free - Sign up +

Business phone number*
13105553890

Email
himiv59877@cmheia.com

Enter a password to sign in to your account.

Create password* Confirm password*
*****| |

We've sent a verification code to himiv59877@cmheia.com. Enter the code to complete sign up.

Verification code*
490869 | [Resend code](#)

I understand that Microsoft may contact me about my trial.
I will receive information, tips, and offers about Power BI, solutions for businesses and organizations, and other Microsoft products and services.
[Privacy Statement](#)

I would like Microsoft to share my information with select partners so I can receive relevant information about their products and services. To learn more, or to unsubscribe at any time, view the [Privacy Statement](#).

By choosing **Next**, you agree to our [terms and conditions](#) and [Privacy Statement](#).

Next Back

3 Confirmation details

Your Privacy Choices Privacy & cookies Terms of use Trademarks © 2024 Microsoft Feedback

Microsoft

You've selected Microsoft Fabric free

1 Let's get you started
2 Create your account
3 Confirmation details

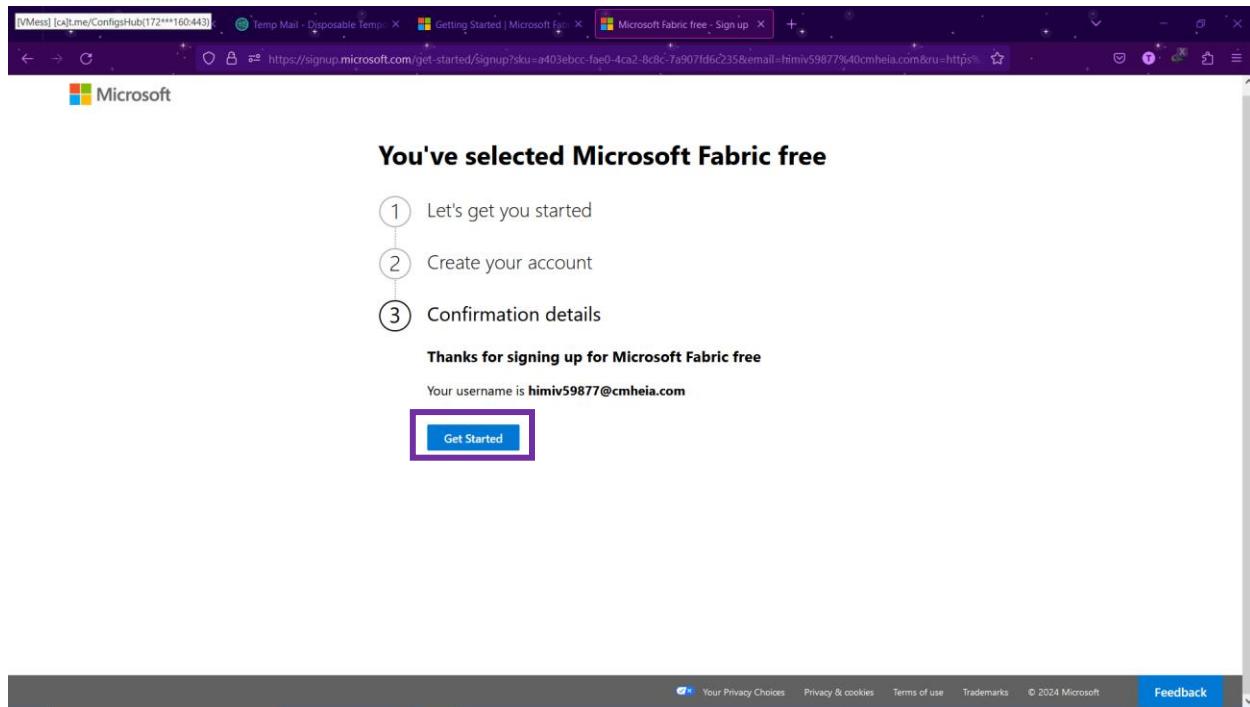
Thanks for signing up for Microsoft Fabric free

Your username is himiv59877@cmheia.com

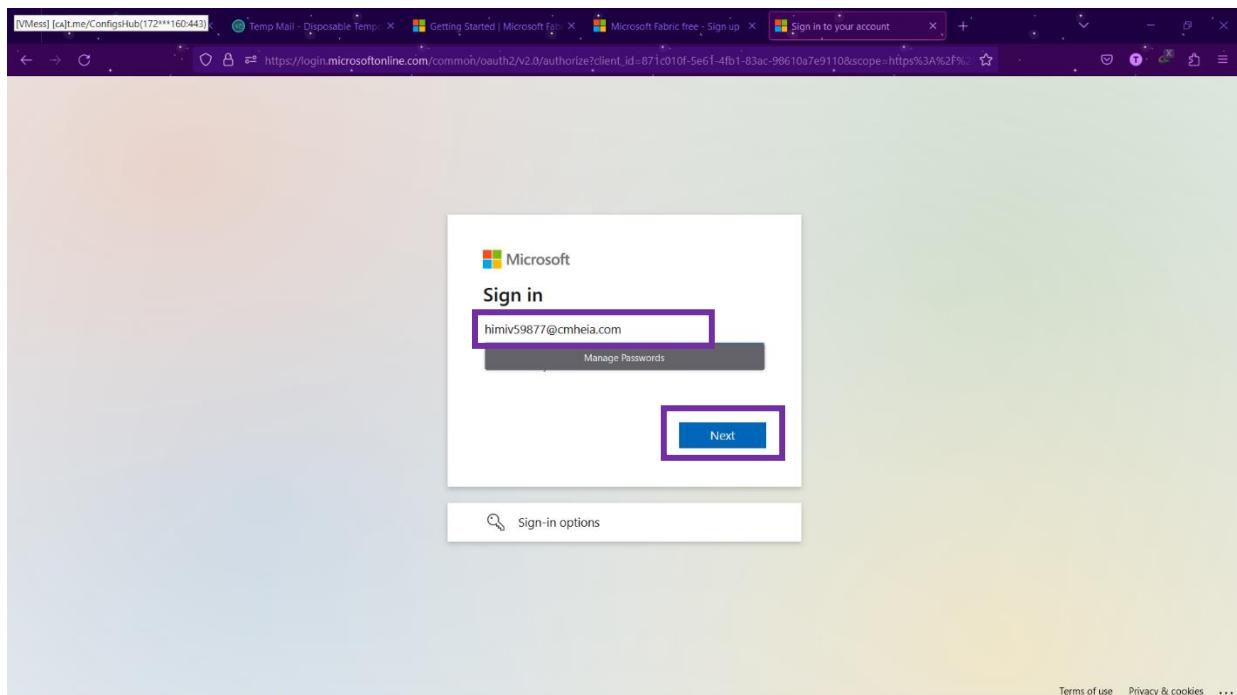
Get Started

Your Privacy Choices Privacy & cookies Terms of use Trademarks © 2024 Microsoft Feedback

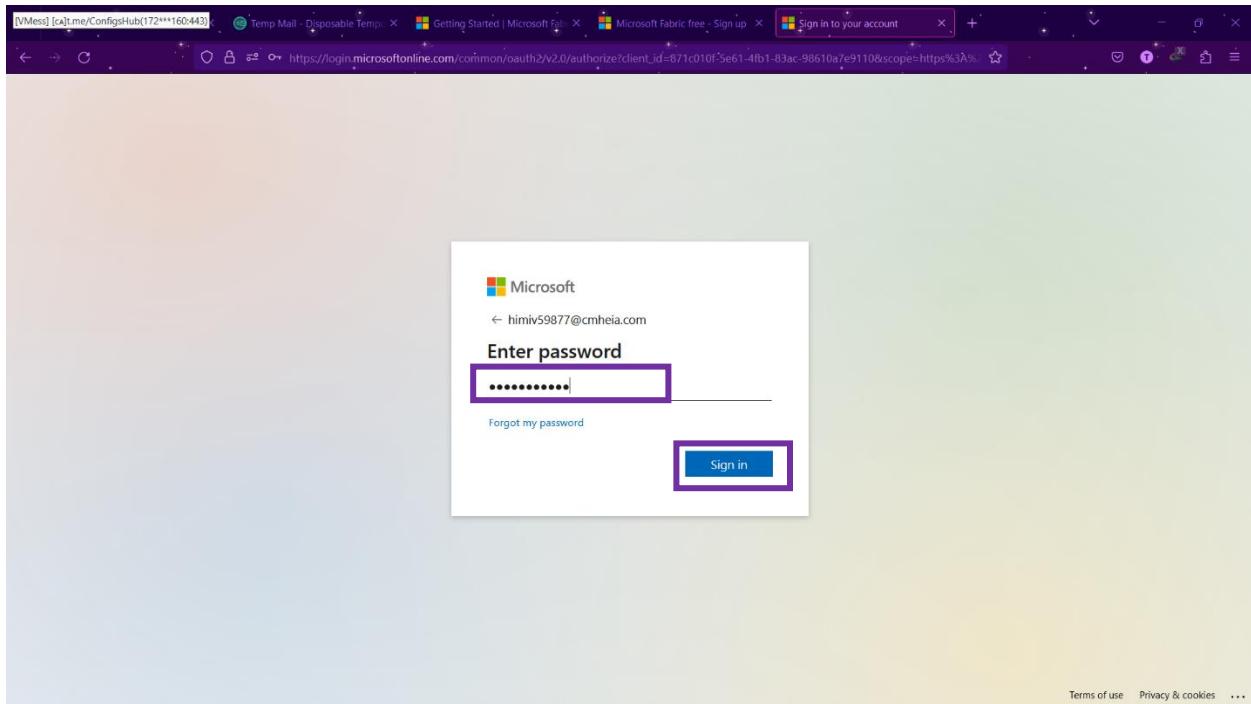
رو میزnim Get started



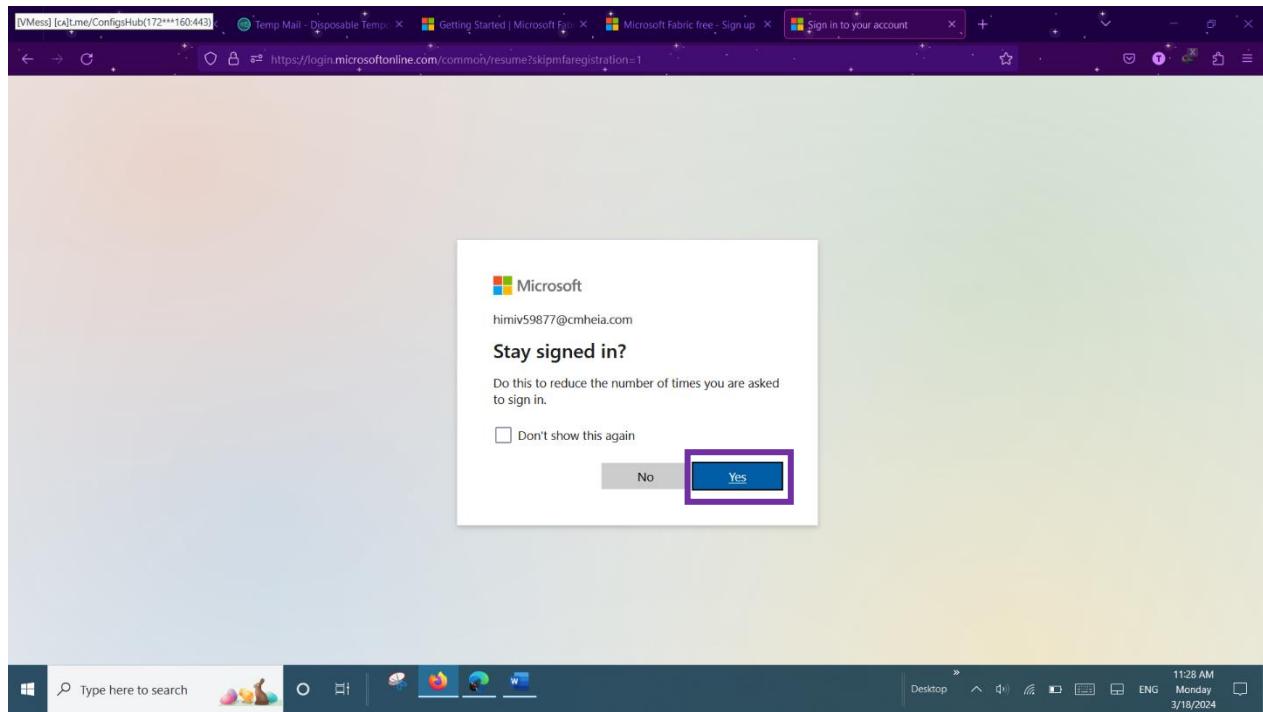
اطلاعات حساب خود را وارد میکنیم. Next میزنیم.



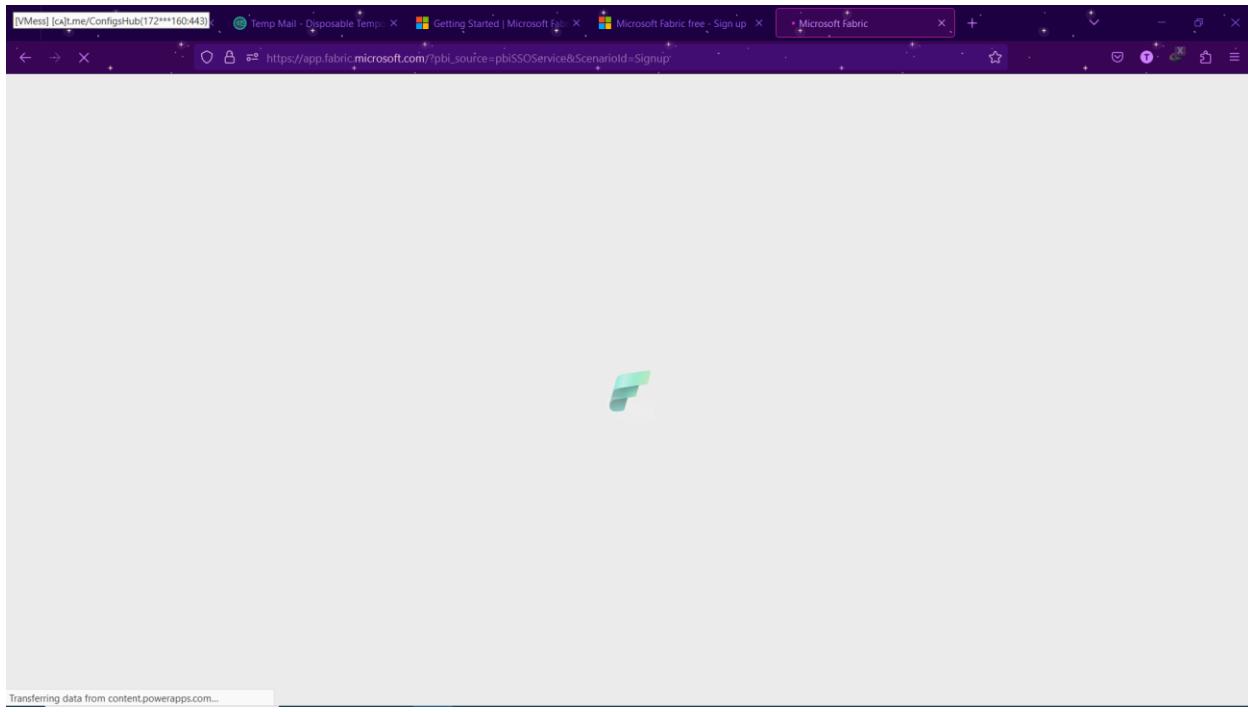
پسورد خود را وارد میکنیم و بعد Sign in



ازمون میپرسه که آیا در این سیستم حسابمون رو نگه داره و بدون گرفتن اطلاعات دوباره ورود کنیم. yes رو میزنیم.



خب ما تازه وارد مایکروسافت فابریک شدیم.



آشنایی با محیط Microsoft Fabric

وقتی برای اولین بار وارد فلایریک میشویم این چیزی است که به نظر می رسد به این 7 تجربه مختلف تقسیم می شود که احتمالاً با بسیاری از آنها آشنا هستید. اما اینها واقعاً به چه معنا هستند؟ اینها شخصیت هایی هستند که تجربیات گروهی از ابزارهای مختلف هستند.

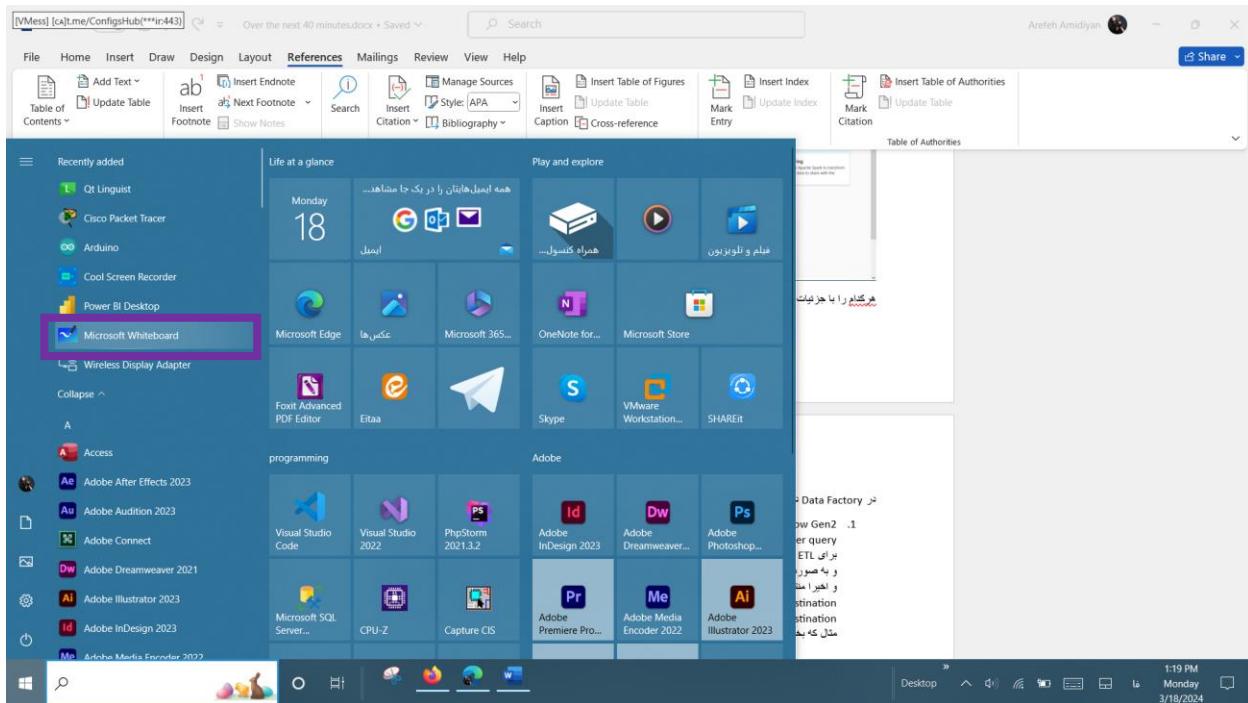
Microsoft Fabric

All your data. In one location. Organize. Collaborate. Create.

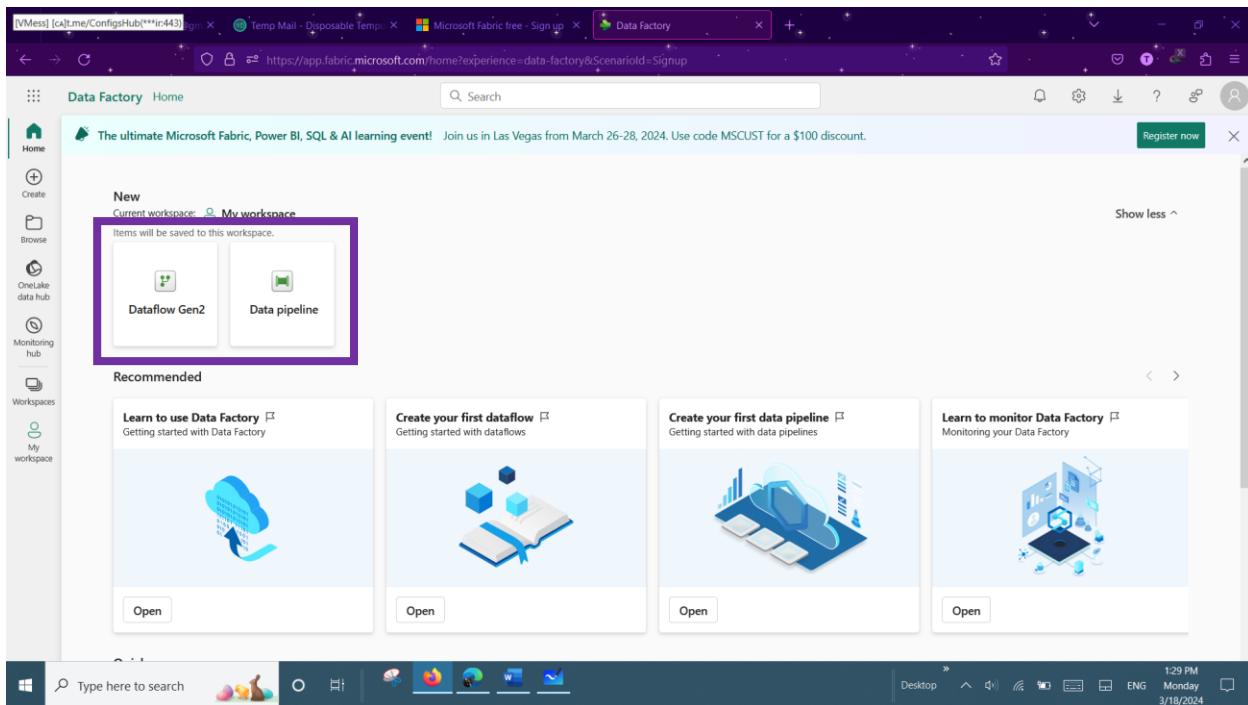
Explore the experience

- Power BI**
Find insights, track progress, and make decisions faster using rich visualizations.
- Data Factory**
Solve the most complex data integration and ETL scenarios with cloud-scale data movement and data transformation services.
- Data Activator**
Monitor data to trigger alerts and automated actions so your organization adapts to changing conditions in real time.
- Synapse Data Engineering**
Create a lakehouse, and use Apache Spark to transform and prepare organizational data to share with the business.
- Synapse Data Science**
Explore your data, and build machine learning models to infuse predictive insights into your analytics solutions and applications.
- Synapse Data Warehouse**
Scale up your insights by storing and analyzing data in a secure, open-data-format SQL warehouse with top performance at PB scale.
- Synapse Real-Time Analytics**
Rapidly ingest, transform, and query any data source and format, from 1 GB to 1 PB, and then visualize and share the insights.

هر کدام را با جزئیات بررسی میکنیم که بینیم با هر کدام چه کاری میتوان انجام داد. در واقع همه با هم مرتبط هستند. در وایت برد مایکروسافت شروع به توضیح داده میکنیم.



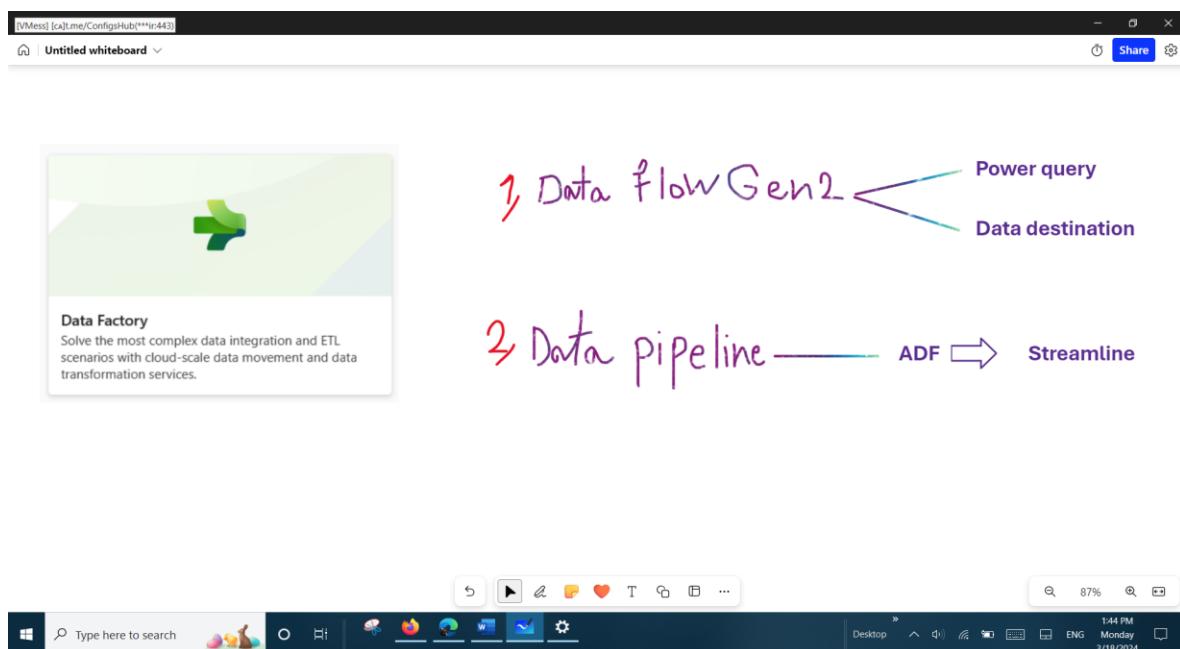
با شروع میکنیم.



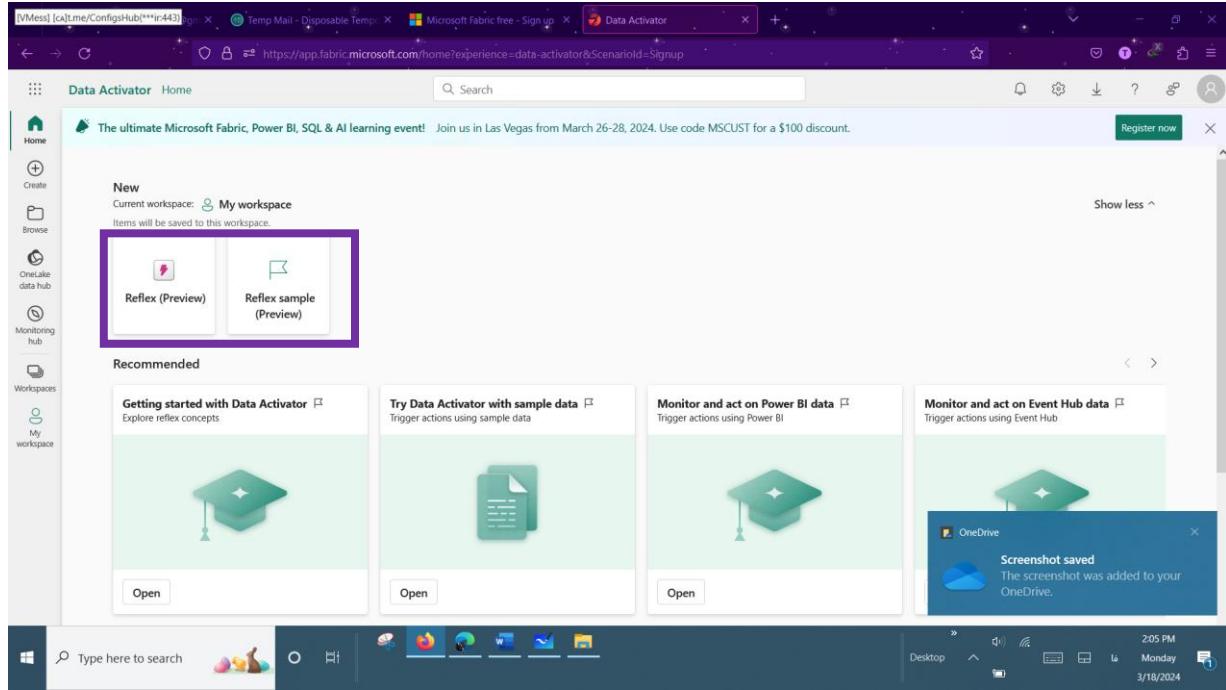
در گزینه دو وجود دارد:

1. Dataflow Gen2: اگر شما کاربر BI Power BI باشید احتمالاً با آن آشنایی دارید و به شما این امکان را می‌دهد Power query بنویسید یا نیازی به نوشتن آن ندارید، اما با رابط کاربری هم شما میتوانید Power query بسازید و اینجا برای ETL است. ما میتوانیم داده‌ها را از رابط‌های مختلف پایگاه‌های داده SQL بگیریم یا هر نوع پایگاه داده و به صورت باقی API HTTP. همه این چیز‌های مختلف آن را به نوعی تغییر می‌دهند. کلیدی که فناوری با Gen2 است و اخیراً منتشر شده است را تغییر می‌دهند و برای فابریک هم ساخت شده است، به شما امکان می‌دهد که Data destination را اضافه کنید تا بتوانیم داده‌ها را از اینجا بیرون بشیم و بتوانیم آن را کمی تغییر دهیم و ما می‌توانیم یک Date destination را تعیین کنیم که می‌تواند ما باشد یا می‌تواند یک Lakehouse باشد به طور مثال که بخشی از مهندسی داده است.

2. دومین عنصر ADF Data pipeline است. اگر ADF Data pipeline استفاده کرده اید این خطوط بسیار شبیه به هم هستند. کمی ساده‌شده‌اند. به این معنی که شما می‌توانید فعالیت‌های کپی داده را انجام دهید، می‌توانید کارهایی مانند control flow logic را انجام دهید اگر سپس statement ها رو switch کنید. به شما این امکان را می‌دهد که اساساً دوباره داده‌ها را از یک مکان بگیرید و برخی از تبدیل‌ها را انجام دهید و سپس آنها را در قسمتی دیگر از فابریک بنویسید.



بعدی Data Activator هست.



یک تجربه بدون کد در مایکروسافت فایریک برای انجام خودکار اقدامات در هنگام شناسایی الگوهای شرایط در تغییر داده است. بدادهای گزارش‌های Power BI و موارد Eventstreams را برای زمانی که داده‌ها به آستانه‌های خاصی می‌رسند یا با الگوهای دیگر مطابقت دارند، نظارت می‌کند. سپس به طور خودکار اقدامات مناسبی مانند هشدار دادن به کاربران یا راه اندازی جریان‌های کاری Power Automate انجام می‌دهد.

به مشتریان این امکان را می‌دهد که یک سیستم عصبی دیجیتال بسازند که در تمام داده‌های آنها، در مقیاس و به موقع عمل می‌کند. کاربران تجاری می‌توانند شرایط کسب‌وکار را در یک تجربه بدون کد توصیف کنند تا اقداماتی مانند ایمیل، اعلان‌های تیم‌ها، جریان‌های Power Automate و فراخوانی به سیستم‌های اقدام شخص ثالث را اجرا کنند. کاربران تجاری می‌توانند نیازهای خود را برآورده کنند و انتکای خود را به تیم‌های داخلی IT و یا توسعه دهنگان کاهش دهند، که هر کدام از این دو اغلب پرهزینه است و مانع چابکی می‌شود. سازمان‌های مشتری برای مدیریت و حفظ رامل‌های هشدار یا نظارت داخلی سفارشی نیازی به تیم توسعه‌دهنده ندارند.

برخی از موارد استفاده رایج عبارتند از:

- هنگامی که فروش در همان فروشگاه کاهش می‌یابد، تبلیغات را اجرا کنید.
- به مدیران فروشگاه‌ها هشدار دهید که مواد غذایی را قبل از فاسد شدن از فریزرهای خراب فروشگاهی مواد غذایی خارج کنند.
- مشتریانی را که تجربه بدی داشتند با ردیابی سفر آنها از طریق برنامه‌ها، وب سایت‌ها و غیره حفظ کنید.
- به شرکت‌های تدارکات کمک کنید محموله‌های گشده را به طور فعال پیدا کنند، وقتی وضعیت بسته برای مدت زمان مشخصی به روزرسانی نمی‌شود، با شروع یک گردش کار تحقیقاتی، محموله‌های گشده را پیدا کنند.
- وقتی مشتریان چهار معوقات می‌شوند، تیم‌های حساب را با محدودیت‌های زمانی یا ارزش سفارشی شده برای هر مشتری، هشدار می‌دهند.

- ردیابی کیفیت خط لوله داده، یا اجرای مجدد کارها یا هشدار در صورت خرابی خطوط لوله یا شناسایی ناهنجاری‌ها.

Core concepts

مفاهیم زیر برای ایجاد و راه اندازی اقدامات و پاسخ‌های خودکار در Data Activator استفاده می‌شود.

Events

Data Activator تمام منابع داده را جریانی از رویدادها در نظر می‌گیرد. یک رویداد، مشاهده‌ای در مورد وضعیت یک شی است، با یک شناسه برای خودشی، یک مهر زمانی و مقادیر فیلدهایی که نظارت می‌کنید. فرکانس جریان‌های رویداد از چندین بار در ثانیه در مورد حسگر‌های IoT تا جریان‌های پراکنده‌تر مانند بسته‌هایی که در داخل و خارج از مکان‌های حمل و نقل اسکن می‌شوند، متفاوت است.

داده‌های مشاهده شده از Power BI نیز به عنوان یک جریان رویداد در نظر گرفته می‌شود. در این مورد، رویدادها مشاهداتی هستند که از داده‌ها در یک زمان‌بندی منظم انجام می‌شوند که معمولاً با فرکانس تازه‌سازی مدل معنایی Power BI شما (که قبلاً به عنوان مجموعه داده شناخته می‌شد) مطابقت دارد. این ممکن است فقط یک بار در روز یا حتی یک بار در هفته اتفاق بیفتد - این فقط یک جریان رویداد به آرامی در حال تغییر است.

Objects

اشیاء تجاری که می‌خواهید نظارت کنید می‌توانند اشیاء فریزرها، وسایل نقلیه، بسته‌های کاربران و غیره یا مفاهیم کمتر ملموس مانند کمپین‌های تبلیغاتی، حساب‌ها، جلسات کاربر باشند. در آیتم reflex خود، با اتصال یک یا چند جریان رویداد، انتخاب یک ستون از شناسه‌شی، و مشخص کردن فیلدهایی که می‌خواهید ویژگی‌های شی را ایجاد کنید، شی را مدل‌سازی می‌کنید.

اصطلاح «نمونه شی» به معنای فریزر/خودرو/بسته خاص و غیره است در حالی که «شیء» معمولاً برای تعریف یا کلاس شی استفاده می‌شود. ما در مورد "جمعیت" صحبت می‌کنیم تا به همه نمونه‌های شی اشاره کنیم.

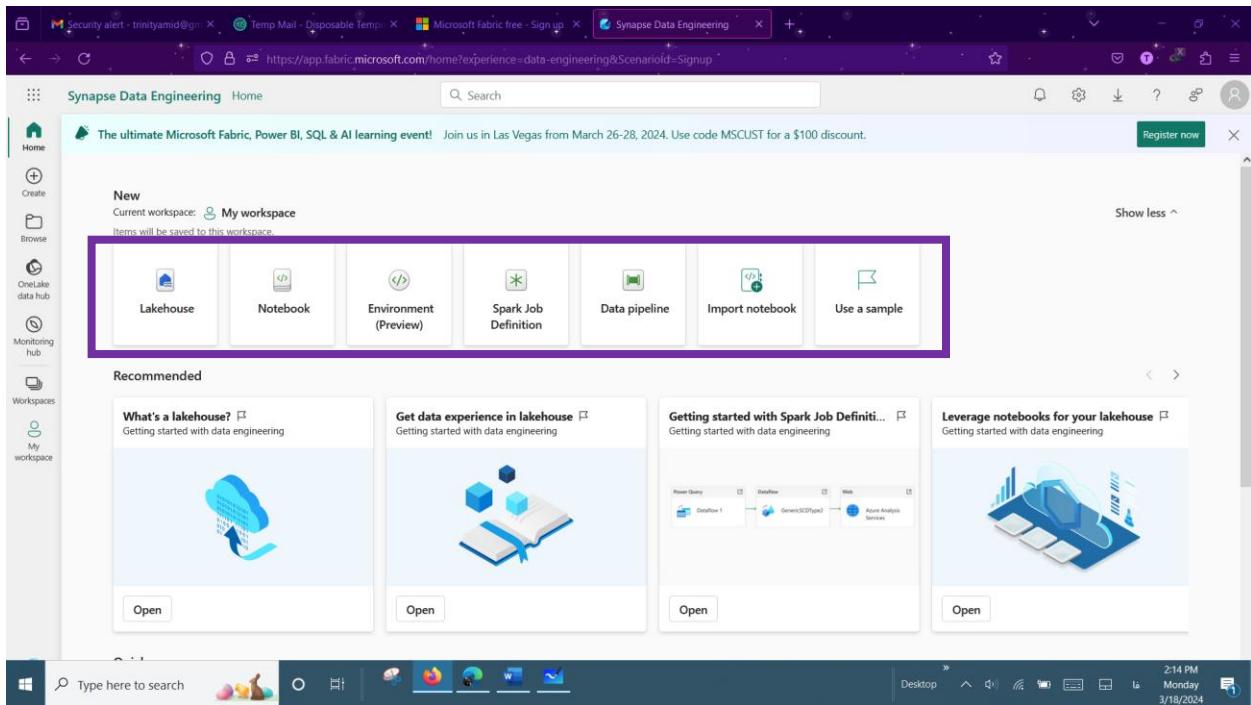
Triggers

ماشه‌ها شرایطی را که می‌خواهید روی اشیاء خود شناسایی کنید و اقداماتی را که می‌خواهید در صورت برآورده شدن آن شرایط انجام دهید، تعریف می‌کنند. هنگامی که یک ماشه فعال می‌شود، همیشه برای یک نمونه شی خاص است. یک ماشه روی یک شی فریزر ممکن است گرم بودن بیش از حد فریزر را تشخیص دهد و یک ایمیل به تکنسین مربوطه ارسال کند.

Properties

ویژگی‌ها زمانی مفید هستند که می‌خواهید از منطق در چندین محرك استفاده مجدد کنید. شما ممکن است یک ویژگی را بر روی یک جسم فریزر تعریف کنید که خوانش دما را در یک دوره یک ساعته صاف کند. سپس می‌توانید از آن مقدار هموار شده در بسیاری از محرك‌ها استفاده کنید.

بعدی هست و این مواردی که نمایش داده شده است در دسترس ما هستند.



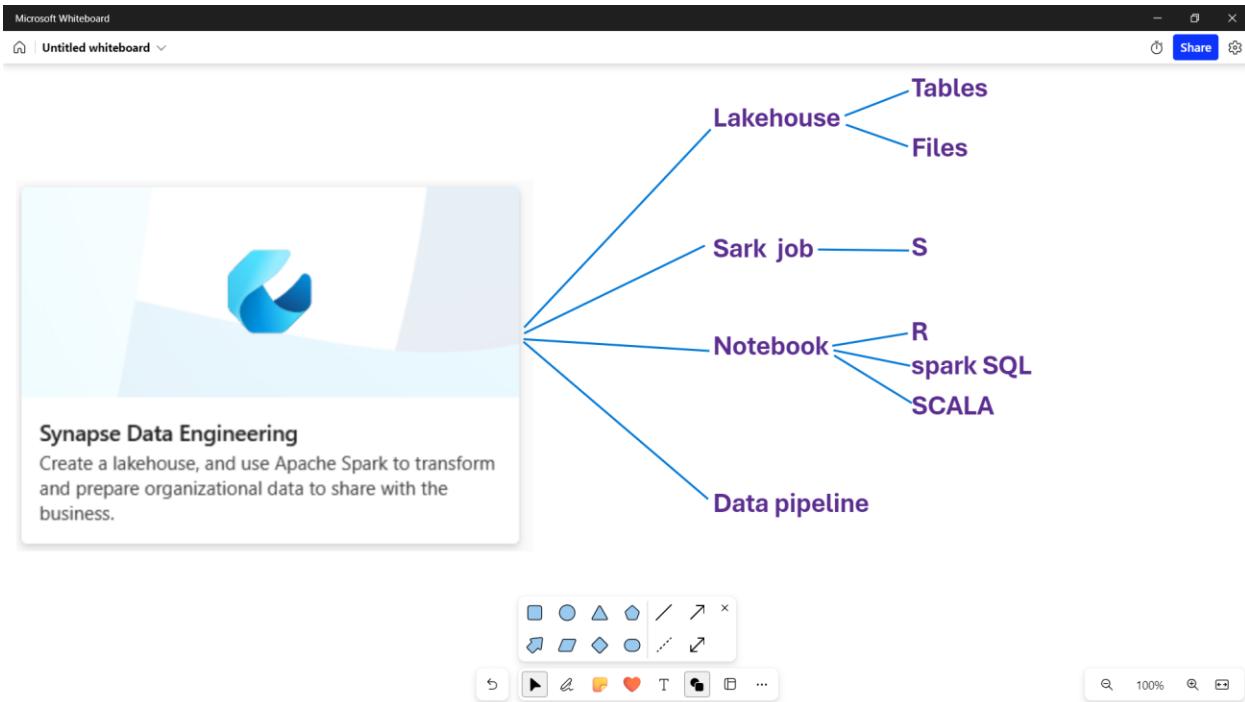
رو داریم که میتوانیم چیزهایی مانند جداول رو در یک **Lakehouse** ایجاد کنیم و همچنین میتوانیم فایل های معمولی را داشته باشیم با فرمات های مختلف.

اما در واقع ما در حال نوشتن **Spark** هستیم موتور زیر این **spark** است، بنابراین به شما امکان می دهد تا پردازش داده های بسیار بزرگی انجام دهید.(مانند peta bytes, tera bytes data)

بدیهی است که میتوانید **Spark jobs** ایجاد کنید که به شما این امکان را می دهد این موارد رو مثل **data transformation, data cleaning,...**

و اگر میخواهید بنویسید در **Notebook** میتوانید بنویسید. و از آنجایی که از روی **Spark engine** ساخته شده، در صورت تمایل می توانید اساسا **Pyspark** را بنویسید همچنین این امکان وجود دارد در **R, spark SQL, SCALA** بنویسید. میتوانید به زبان های مختلف بنویسید. از دیدگاه مهندسی داده این میتواند ساخت اسکریپت ها یا نوت بوک هایی باشد که داده ها رو تمیز می کنند که نوعی تبدیل انجام می دهند اگر قبل از یک **data scientist** کار میکنید و میخواهید به آنها کمک کنید.

آخرین نکته **Data pipeline** است که کپی میشود از **Data factory experience** و فقط به ما کمک می کند تا داده ها را به **Lakehouse** برسانیم. بنابراین اگر داده ها را در یک **data house** دارید، می توانید **pipeline** ای بسازید که داده ها را میخواند، تبدیل می کند و آنها را در **Lakehouse** تان ذخیره می کند تا برای شما آماده شود. از یک (**R, spark SQL, SCALA**) **notebook** بردارید و تغییرات بیشتری را در هر یک از این زبان ها انجام دهید.



جنبه مهندسی داده در سطح بسیار بالایی است بنابراین نگاهی به کارهایی که میشه در بخش Data Science انجام داد میکنیم.

ما در این بخش 3 ابزار اصلی در اختیار داریم.

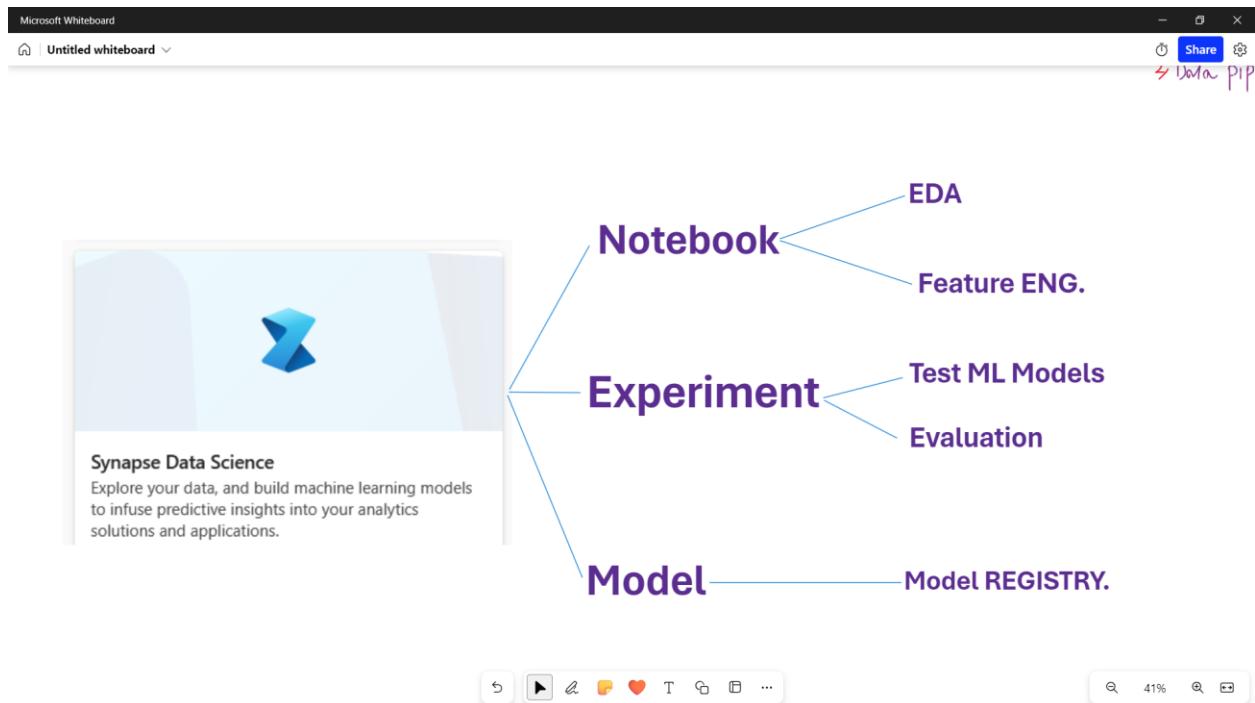
ما یک Data scientist داریم که مشابه Notebook هست که به ما امکان می دهد به عنوان EDA را در یک notebook انجام دهیم یا برخی از Features Eng.

اساسا در یک notebook شما میتوانید چیزهای جدیدی را امتحان کنید، داده ها را کاوش کنید یا هرکاری که یک Data scientist ممکن هست بخواهد انجام دهد.

وقتی کمی داده ها را کاوش کردید، ممکن هست که بخواهید یک مدل یادگیری ماشین بسازید و آن را در برابر برخی معیارهای ارزیابی آزمایش کنید. اکنون ممکن است چندین مدل داشته باشید که ممکن است بخواهید آزمایش کنید که آیا این مدل جنگل تصادفی بهتر از این مدل رگرسیون لجستیک است یا خیر و برای هر مدل ممکن است پارامترهای هایپر مختلفی داشته باشید که می خواهید آنها را تست کنید، بخش آزمایش همانطور که از نام آن پیداست برای آزمایش مدل های ML است براساس معیارهای ارزیابی بسته به هدف و مسئله ای که میخواهید با این آزمایش حل کنید.

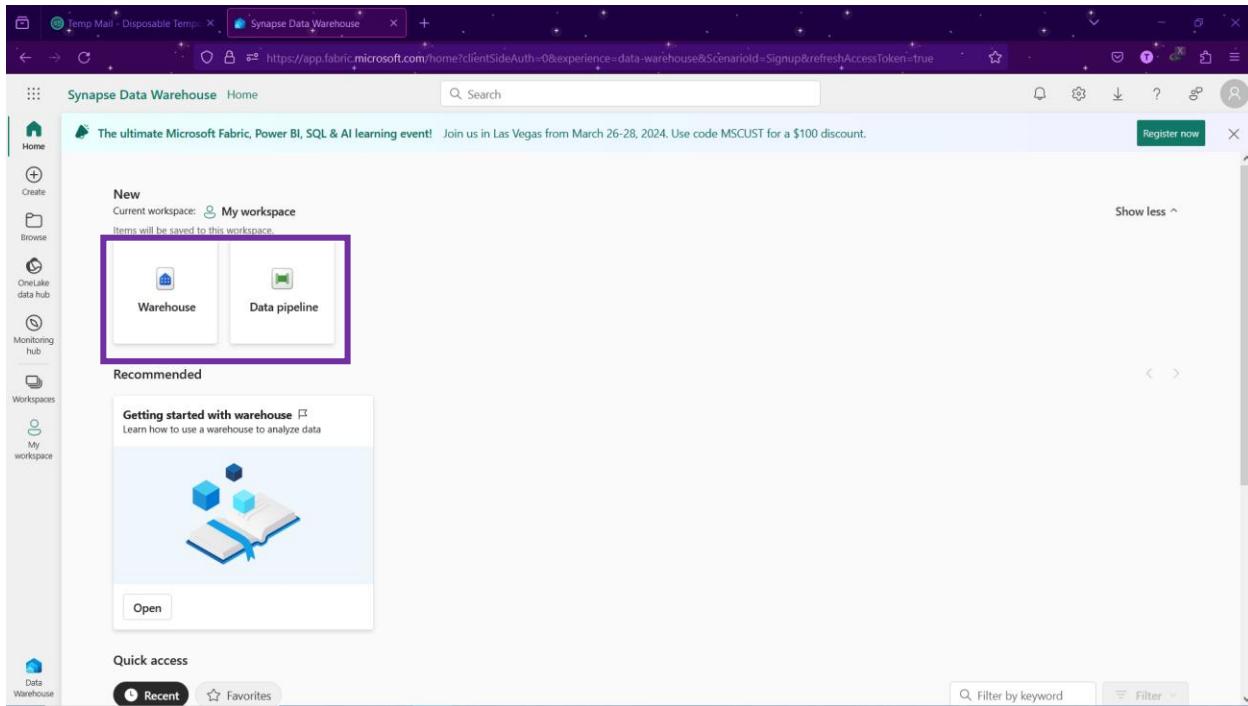
بعد از اینکه مدلی را که دوست دارید پیدا کردید که در برابر برخی معیارهای ارزیابی عملکرد خوبی دارد و شما مجموعه ای از وزن ها را دارید، سپس شما می توانید آن را به عنوان یک مدل ذخیره شده منتشر کنید. در model tab این مفهوم از یک model registry دارید. بنابراین اگر می خواهید یک مدل آموزش دیده را با شخصی در سازمانتان به اشتراک بگذارید در اینجا قرار است این کار را انجام دهید.

این مدل ها می توانند برای مواردی مانند استنتاج استفاده شوند، بنابراین می توانید در این بخش پیش بینی هایی که از مدل ماشین لرنینگ خود به Lakehouse بنویسید تا شاید یک مورد استفاده باشه که میتوانید در آنجا استفاده کنید.



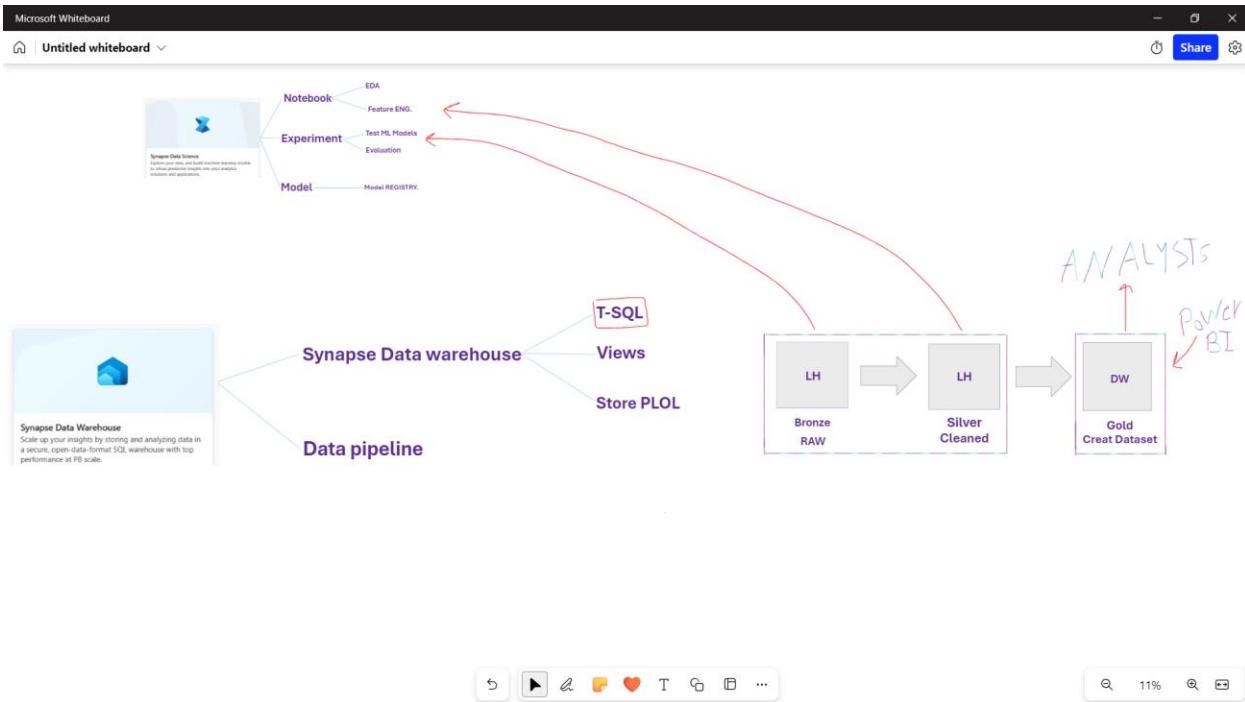
قسمت بعدی Data Warehouse هست.

در warehouse persona بدبیهی است که ما قادر خواهیم بود که یک data warehouse, data pipeline ایجاد کنیم.



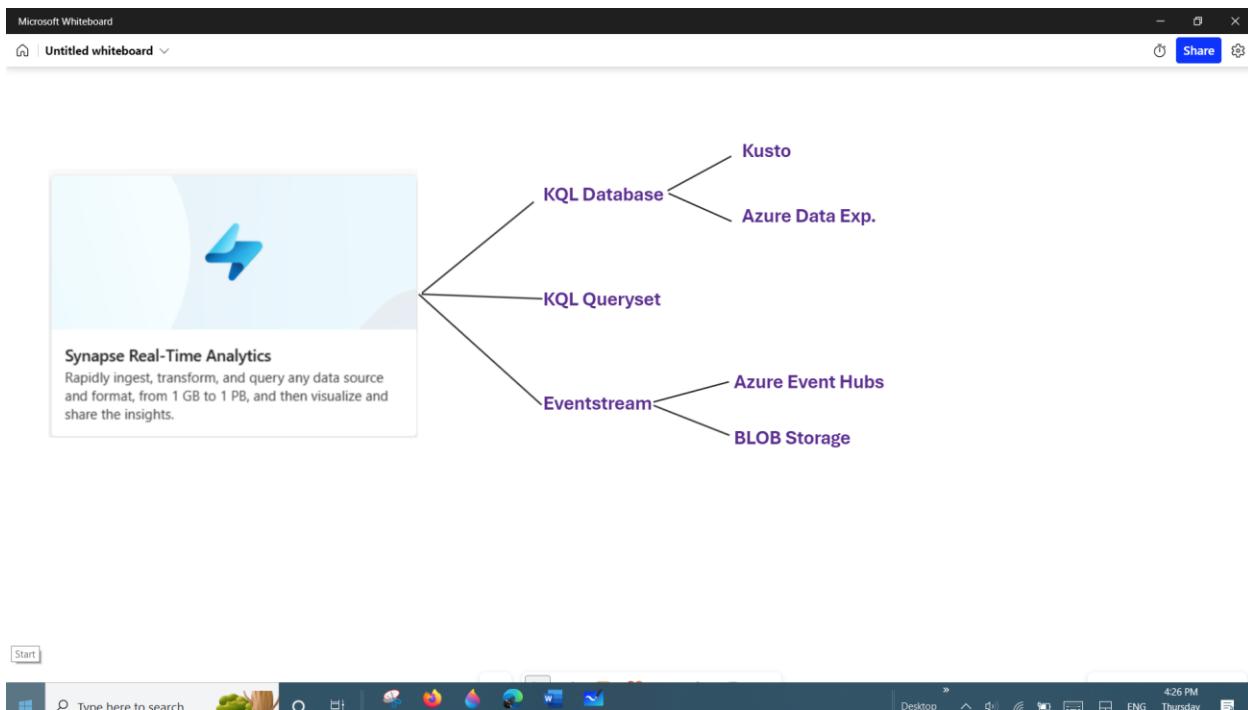
خطوط data warehouse برای دریافت داده ها به data warehouse است و data pipeline به ما امکان می دهد تا T-SQL queries بنویسیم مانند قبل که شما عادت کرده اید در حال حاضر هیچ مفهومی از چیزهایی مانند dedicated pause یا هر چیزی که با نوعی از SQL چیزها همراه است وجود ندارد. اکنون به سادگی data warehouse است. بنابراین در اینجا شما می توانید دیدگاه های خود را بنویسید، رویه ها را رو ذخیره کنید یا همه چیزهای خوب را در حال حاضر در آنجا بنویسید که چگونه این ممکن است در یک سازمان تطبیق داده شود. اگر شما نوعی معماری مدلایون را هدف قرار می دهید ممکن است bronze تان را در اینجا به عنوان Lakehouse داشته باشید که data raw شما است. پس از آن ممکن است Lakehouse دیگری برای silver خود داشته باشید. این خانه به نوعی تمیز می شود، شاید چند notebook ساخته باشید که این داده ها را تمیز می کند. و سپس در انتهای pipeline ممکن است data warehouse خود را داشته باشید که gold شماست. و این برای dataset های انتخاب شده است که احتمالاً توسط Power BI شما مورد بررسی قرار می گیرد. این دو LH احتمالاً بیشتر برای data engineering, data scientist شما جایی است که می خواهید بهترین dataset را برای تحلیلگران خود در سازمان خود تنظیم کنید که dataset های خوب و تمیزی به آنها می دهید تا بتوانند به Power BI خود وصل شوند و مستقیماً با Power BI ادغام می شود و می شود به آسانی مستقیماً با warehouse report را ایجاد کرد.

نکته مهمی که در اینجا باید به آن اشاره شود این است که این موتور بر روی موتور TSQL ساخته شده است و موتورهای زیادی در اینجا وجود دارد که در ادامه جزئیات بیشتری را در مورد نووه ادغام آنها توضیح خواهیم داد.



آخرین مورد Real-Time Analytics هست که ۳ نوع ویژگی مثبت داریم که می توانیم سه منبع ایجاد کنیم و خواهد دید که در بالای زبان جستجوی KQL Kusto ساخته شده است که از Azure data explorer می آید.

دوباره اینجا می توانید Real-Time Data ها را تجزیه و تحلیل و ذخیره کنید تا بتوانید چیزهایی مانند جریان رویداد ایجاد کنید، برای ذخیره سازی طولانی مدت داده های جریانی، فضای ذخیره سازی BLOB را نیز جذب کنید. بنابراین هنوز هم ارتباطی با Azure وجود دارد که ممکن است فکر کنید همه ای این 7 موردی که توضیح داده شد به عنوان یک محصول SAS در دسترس هستند، اگر می خواهید در Real-Time انجام دهید، قطعاً به Azure نیاز دارید. هنوز هم می دانید که از Registering IOT Streams استفاده می کنید که Azure Event House را ثبت می کند.



بنابراین آخرین چیزی که باید به آن توجه کنید Power BI هست. اساساً مانند قبل است اما به طور کامل در این راه حل ادغام شده است. اکنون دوستان زیادی دارد، Data science دارد و همه ای اینها در یک work space قرار دارند.

The screenshot shows the Power BI Home page. On the left, there's a sidebar with icons for Home, Create, Browse, OneLake data hub, Apps, Metrics, Monitoring hub, Learn, Workspaces, and My workspace. The main area features a "Recommended" section with four cards:

- Explore basic Power BI concepts**: Getting started with Power BI. Includes a small icon of a bar chart and a link to "Open".
- Explore the 100 most useful productivity tips**: Explore this data story. Includes a thumbnail of a productivity tips card and a link to "Open".
- Cancer statistics in the USA**: Explore this data story. Includes a map of the USA with cancer data and a link to "Open".
- Intro—What is Power BI?**: Getting started with Power BI. Includes a thumbnail of a dashboard and a link to "Open".

At the bottom of the page are buttons for "Recent", "Favorites", and "My apps", along with search and filter tools. The taskbar at the bottom shows various pinned apps like Mail, Photos, and Edge.



Lakehouse

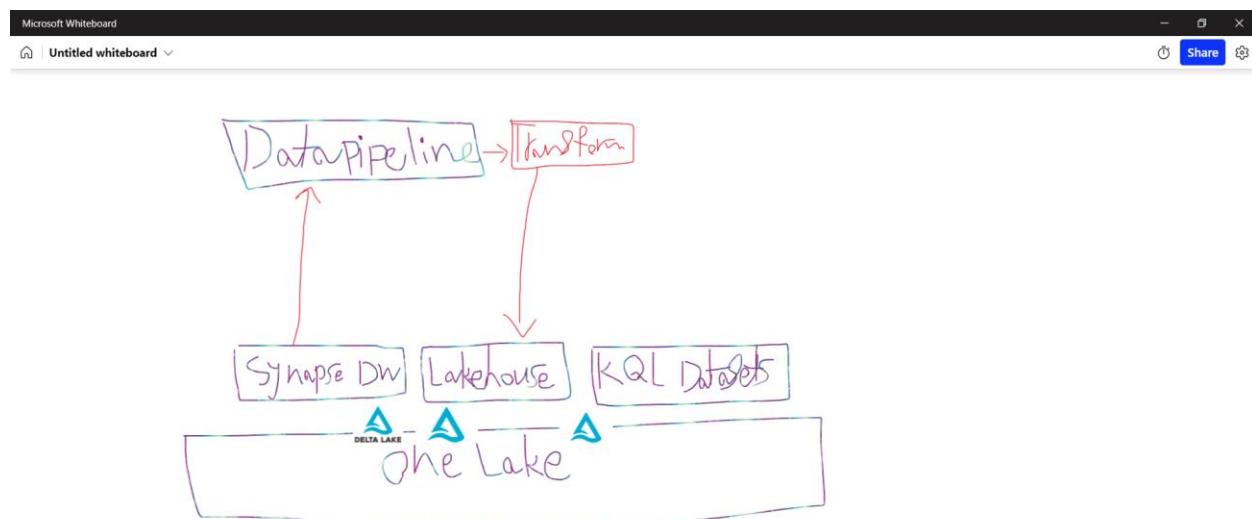
یک مفهوم بسیار مهم و مبتکرانه ترین چیز در فابریک هست که ممکن است درباره‌ی آن شنیده باشید. بنابراین روشی که آنها این مایکروسافت را می‌فروشند این است که شما یک درایو برای اسناد خود دارید و اکنون یک Lake برای داده‌های شما وجود دارد. دو چیز وجود دارد که باید به خاطر بسپارید:

1. فقط یک Lake برای یک سازمان وجود دارد.
2. و شما نباید بدون آن data lake های خودتان را کپی کنید.

اگر از Azure data Factory هنگامی که دیتاها را از lake به اینجا کپی کنید و اگر هر روز این کار را انجام می‌دهید، این واقعاً گران است و حجم زیادی از داده‌ها هدر می‌رود. بنابراین با یک Lake همه داده‌های شما در آن قرار می‌گیرد.

این یک پیوند است و همه این انواع مختلف پایگاه داده‌ای که بررسی کردیم، warehouse، lakehouse، KQL است و همه اینها در یک Lakehouse قرار می‌گیرند و در قالب دلتا قرار دارند. این یک قالب ذخیره‌سازی open source است به دلیل اینکه در قالب یکسانی نوشته شده و خوانده می‌شود به ما امکان می‌دهد واقعاً در حال حاضر کارهای بسیار جالبی انجام دهیم. مانند یک data pipeline که دیتا را از data warehouse بخوانیم و با خواندن Pipeline خود آن را تبدیل می‌کنیم و انتقال می‌دهیم. سپس ممکن است بخواهیم آن را در یک Delta بنویسیم و این امکان پذیر است زیرا همه چیز به نوعی در این قالب Delta ذخیره می‌شود. interfaces، engines وجود دارند که همه با این قالب Delta وصل می‌شوند. یکی از قدرتمندترین چیزهای در مورد فابریک است برای درک.

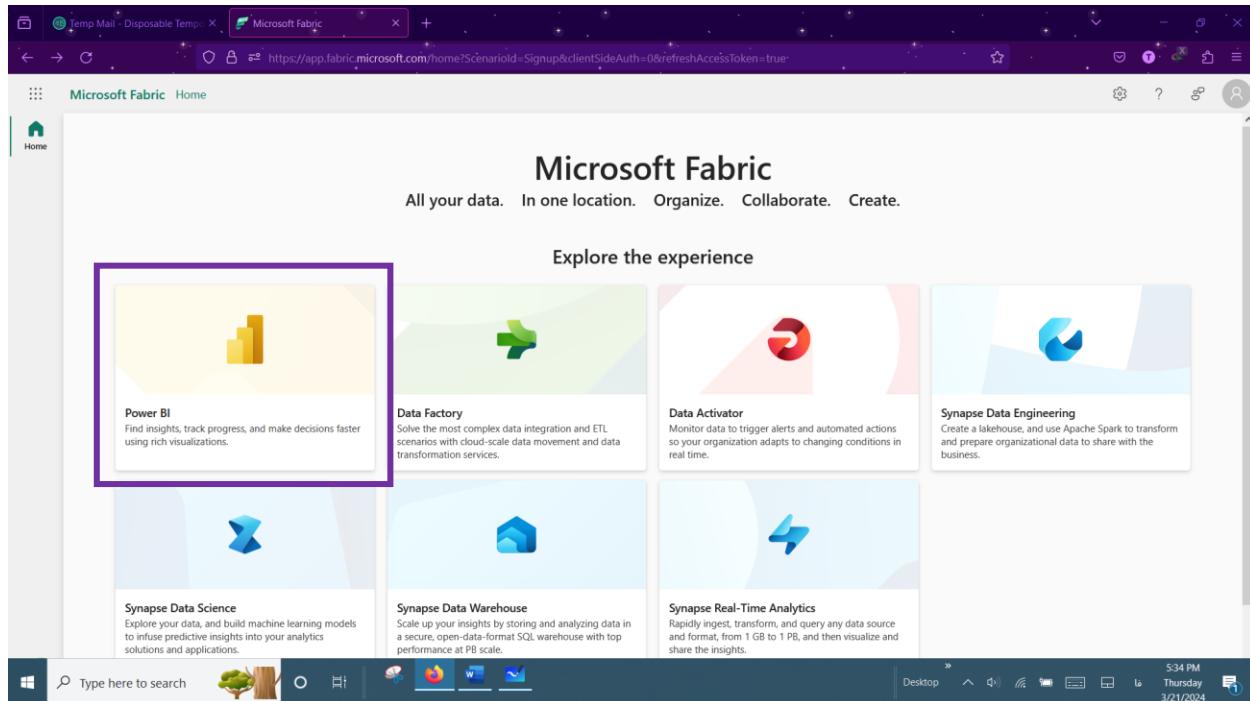
نکته کلیدی دیگری که در اینجا هست این است که به مشتریان کمک می‌کند تا احساس امنیت بیشتری داشته باشند همه داده‌های آنها در open format است، بنابراین اگر در مدت 10 سال آنها بخواهند ارسال کنند و به محصول دیگری بروند، داده‌ها دیگر در قالب‌های اختصاصی ذخیره نمی‌شوند، اما اکنون در قالب‌های open source ذخیره می‌شوند، این بدان معنا نیست که آنها کار را برای شما آسان یا ارزان می‌کنند اما این یک option است.



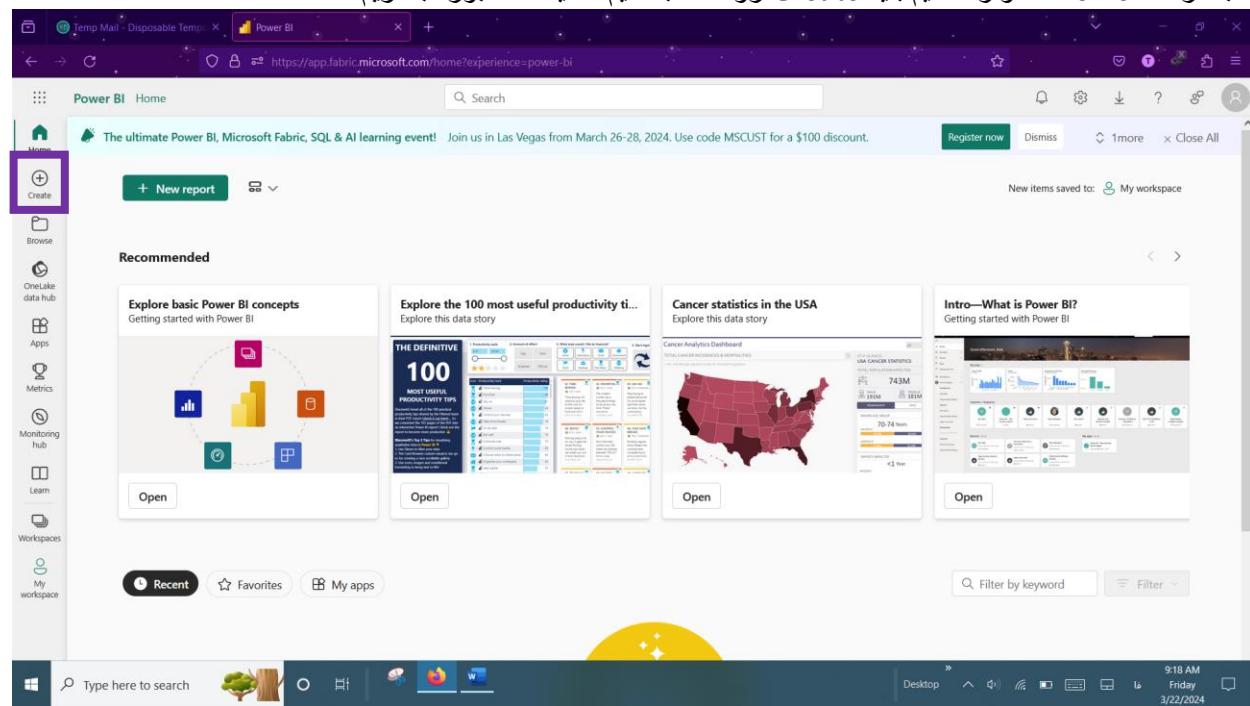
یک مفهومی داریم به اسم **Workhouse** همه‌ی کارهای ما در اونجاست که هر leke چندتا workspace رو درون خودش می‌توانه داشته باشه و به اون چند **DOMAIN**، **workspace** می‌گویند. ممکن است دامنه دیگری برای سازمان خود داشته باشد.

پیاده سازی پروژه در Microsoft Fabric

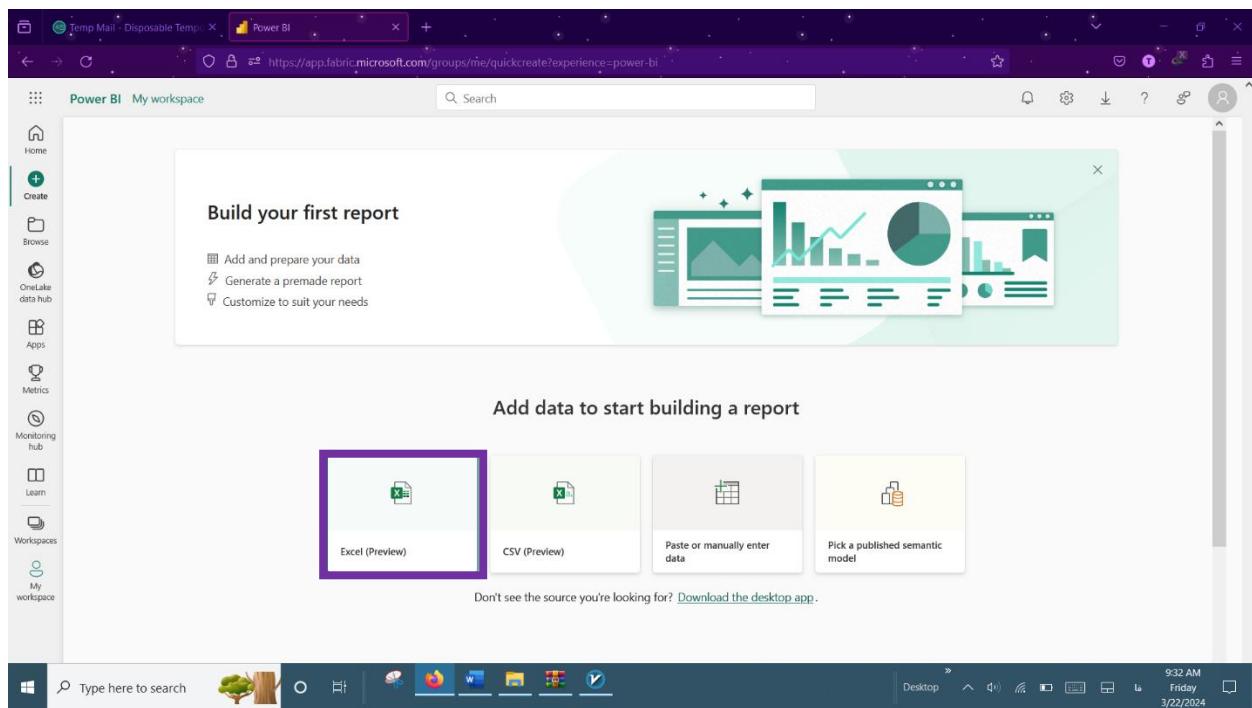
خب ما میخواهیم به پیاده سازی در فابریک بپردازیم. خب کدوم یکی از این 7 تا رو باید انتخاب کنیم و به پیاده سازی داشبورد بپردازیم. پس اینکه تک تک موارد رو مورد بررسی قرار دادیم متوجه میشویم که علاوه بر اینکه قسمت های دیگر فابریک به جز power bi برای ما محدود است و نمی توانیم از آن استفاده کنیم، فقط در power Bi باید به پیاده سازی بپردازیم، بقیه قسمت ها امکان ساخت داشبورد را ندارند.



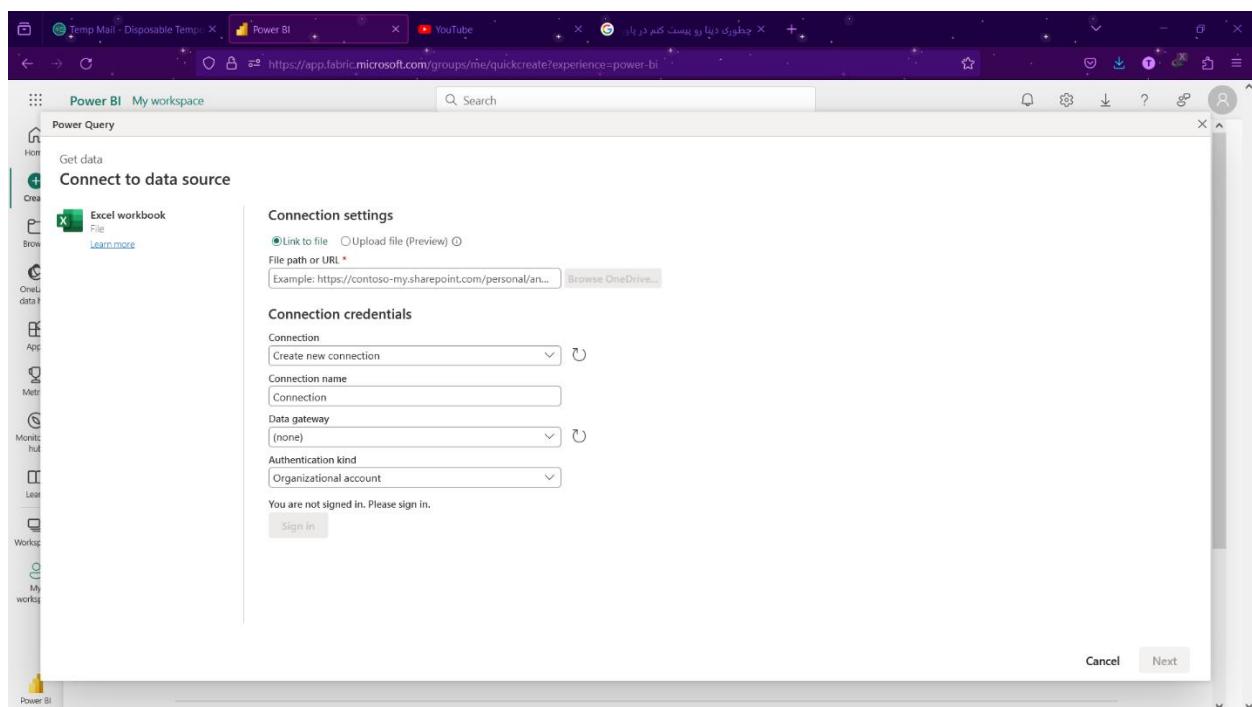
خب در Power BI که وارد شدیم باید Create رو انتخاب کنیم که یک داشبورد بسازیم.



در این قسمت با قسمت های مختلفی روبرو هستیم.

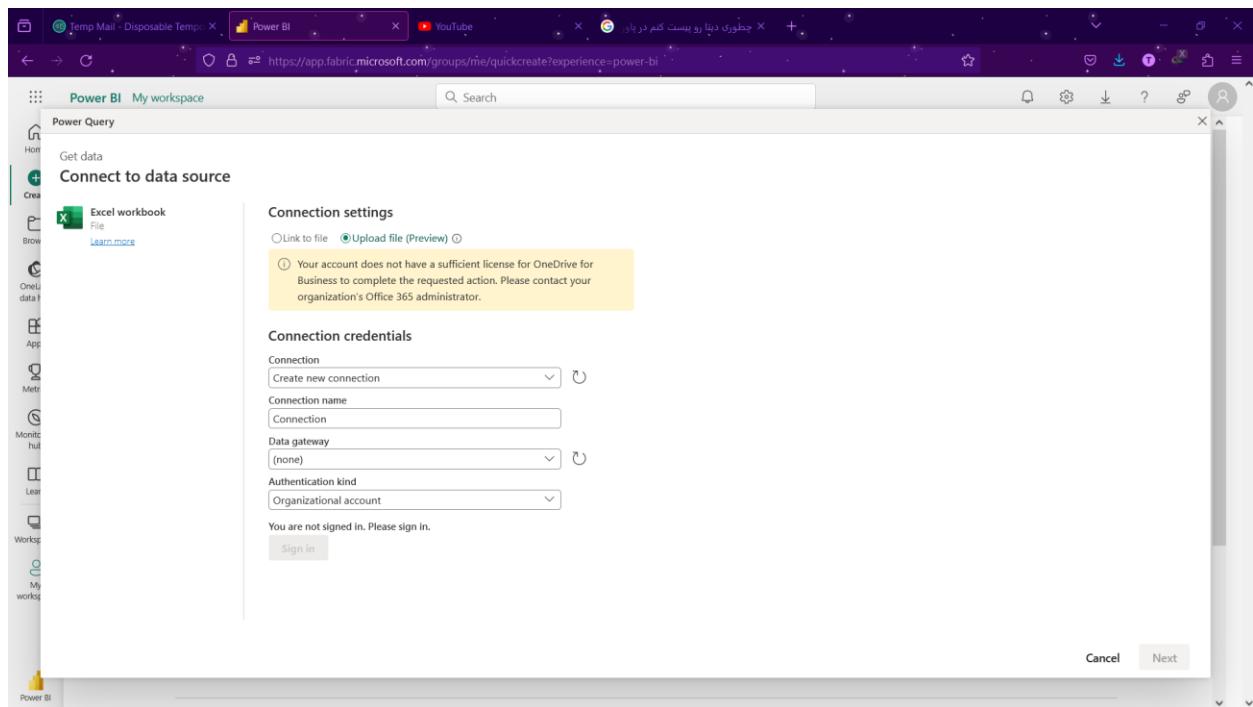


روی Excel(preview) میزنیم و این صفحه برای ما باز میشود.

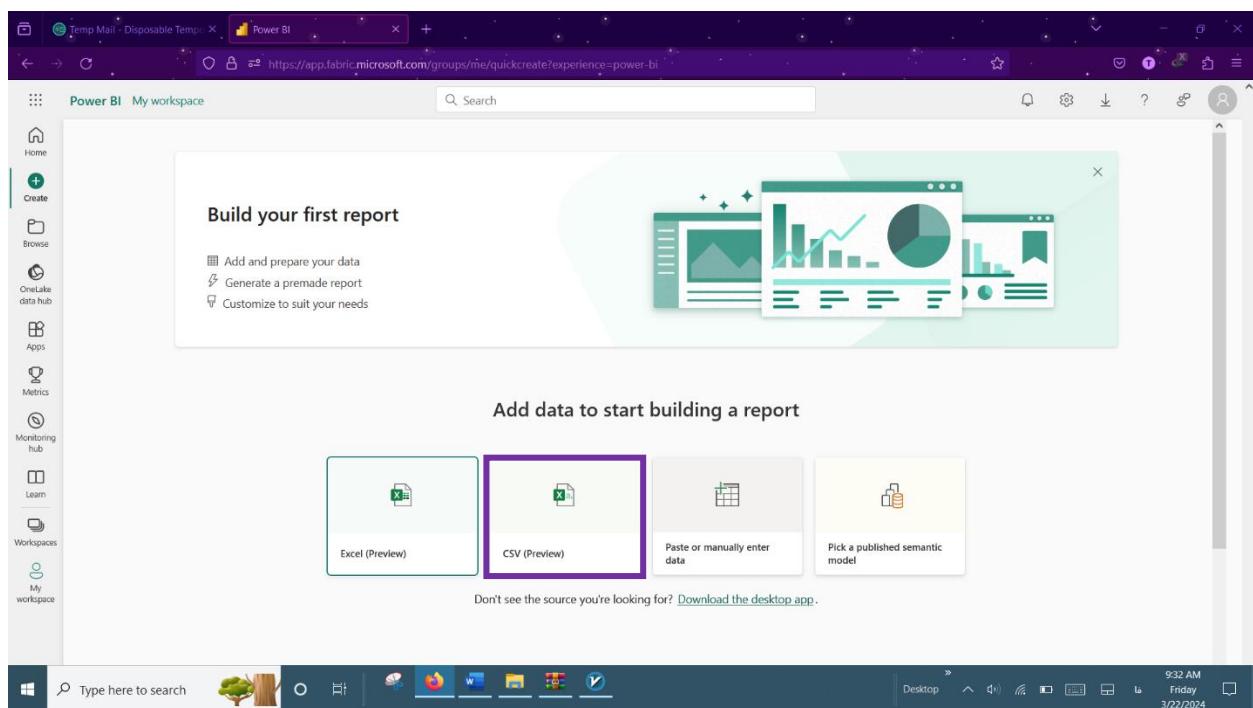


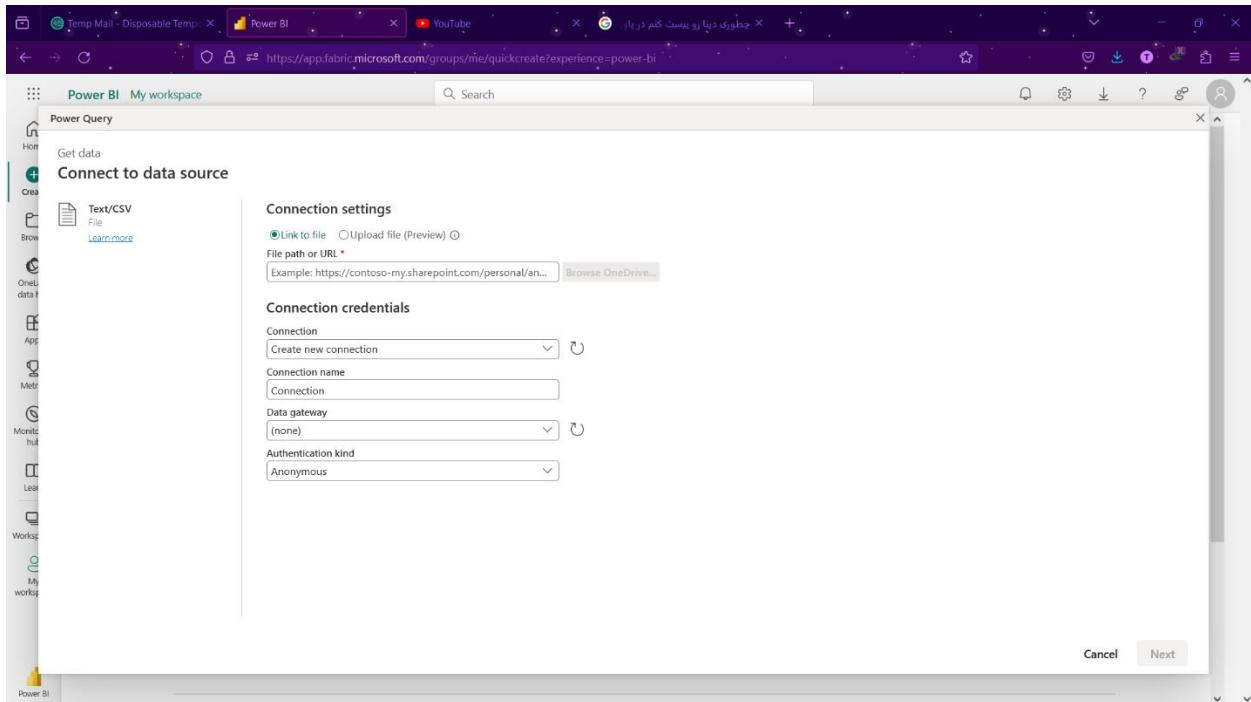
در اینجا میتوان دیتابیس ساخت به صورت لینک به فایل یا آپلود فایل.

روی آپلود فایل که میزnim برآمده میاره که لاینسنس ما محدود هست.

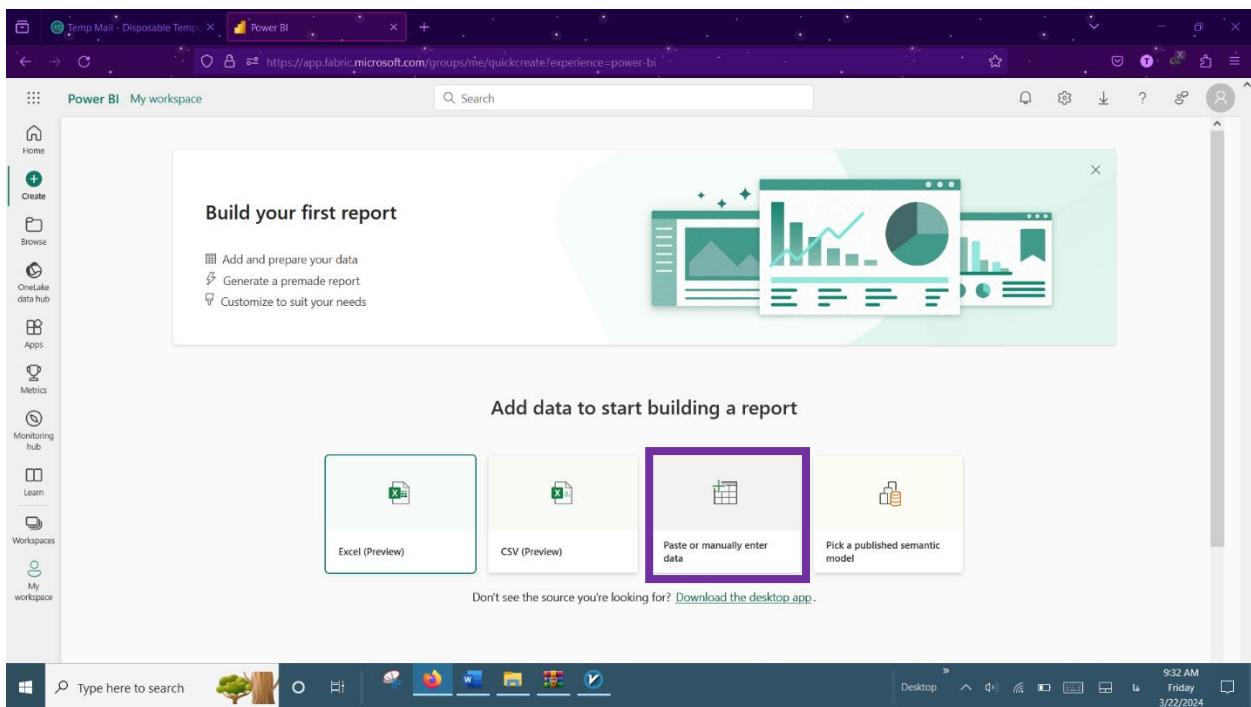


خب در قسمت بعدی که CSV(preview) بود که میزnim، میاره که به یک فایل Text/CSV لینک شود. که چون فرمت فایل ما اکسل هست به کار نمی آید.

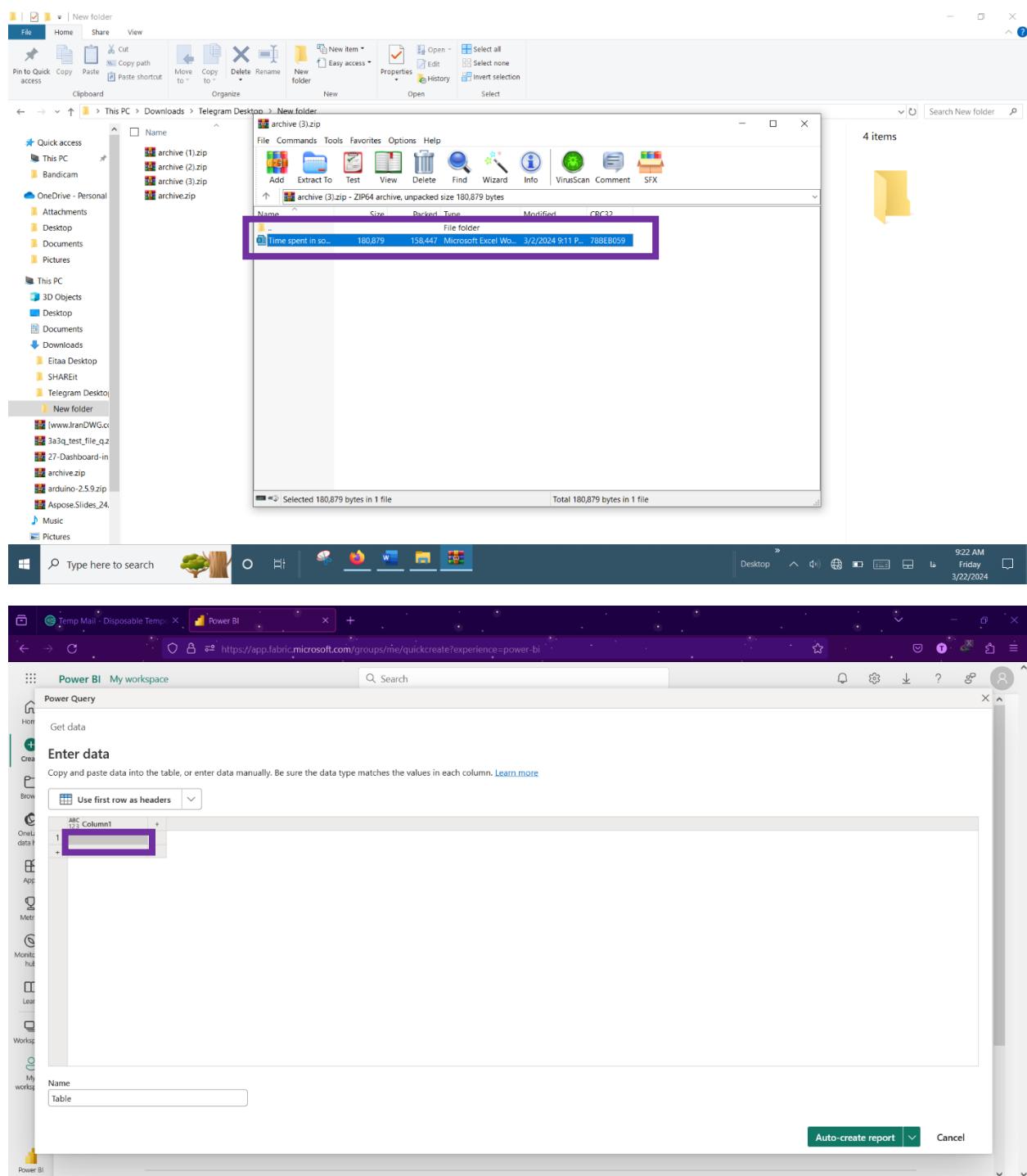




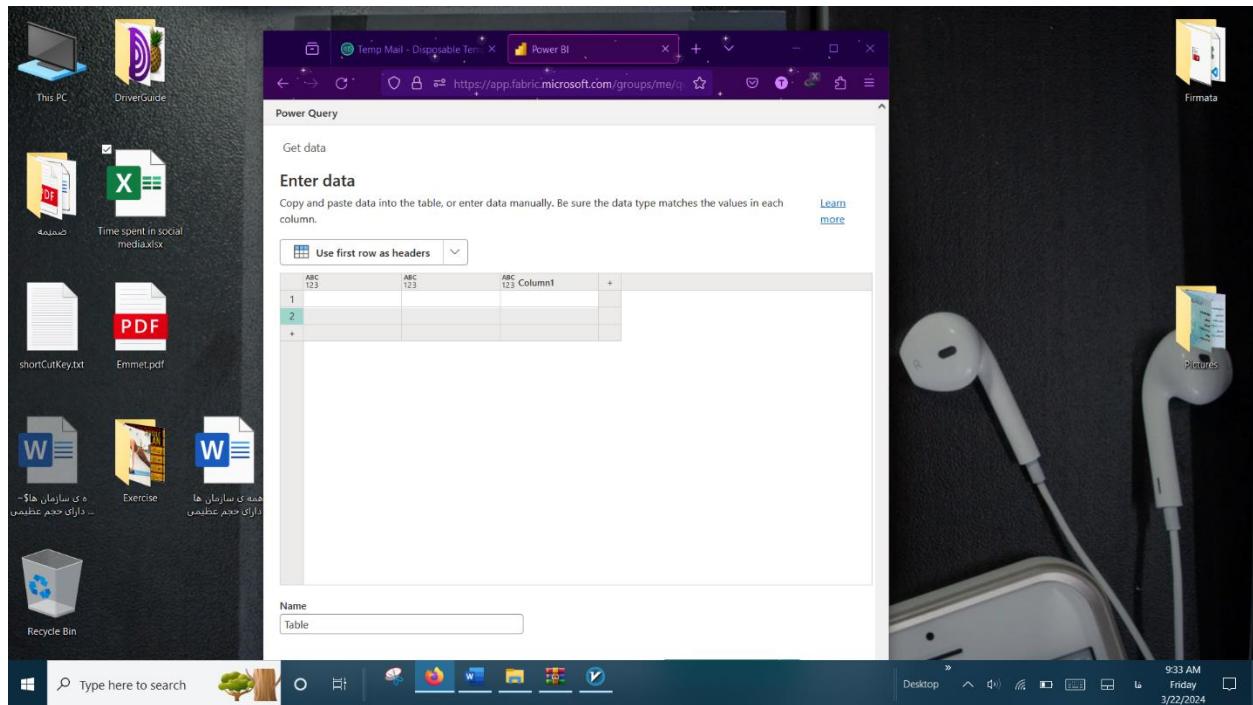
رو انتخاب می کنیم که فایل رو Paste or manually enter data



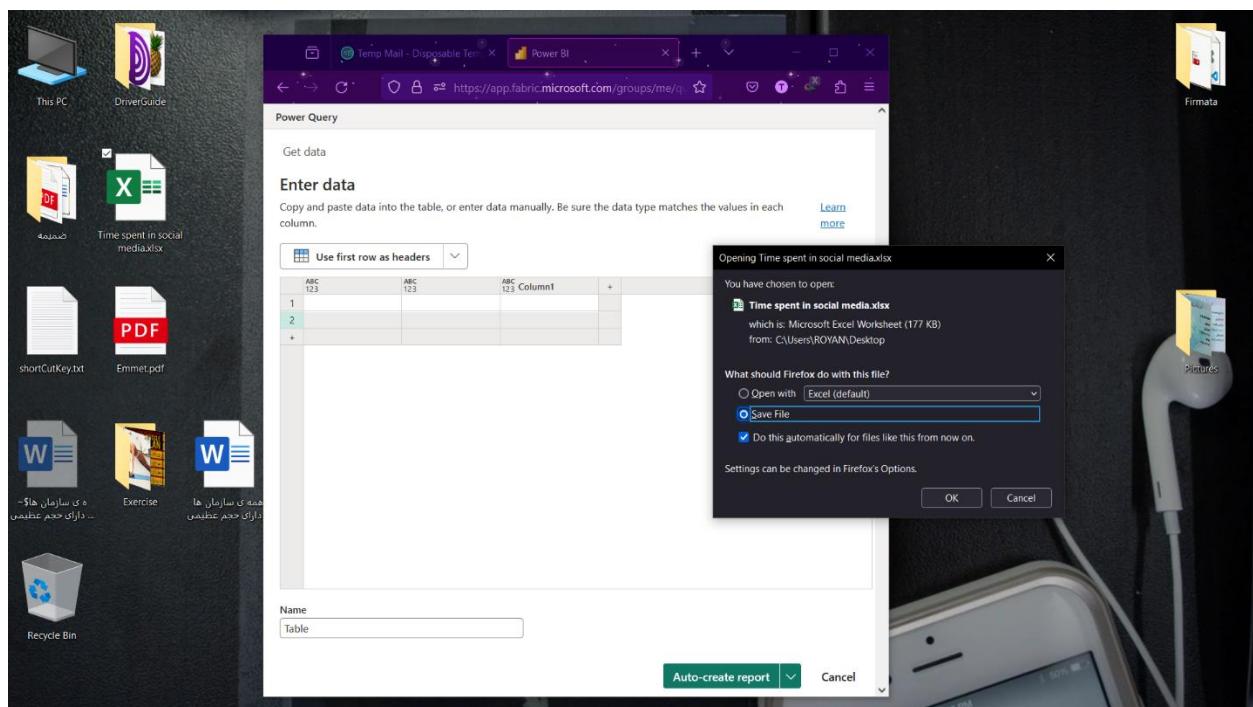
فایل رو از سایت kaggle دانلود میکنیم و در اینجا انتخاب میکنیم بعد کپی میکنیم.



میبینیم که paste نشد. فایل رو extract میکنیم. بعد drag و روی صفحه میاریم.



روی حالت دانلود فایل میرود. که کار ما را انجام نمی دهد.



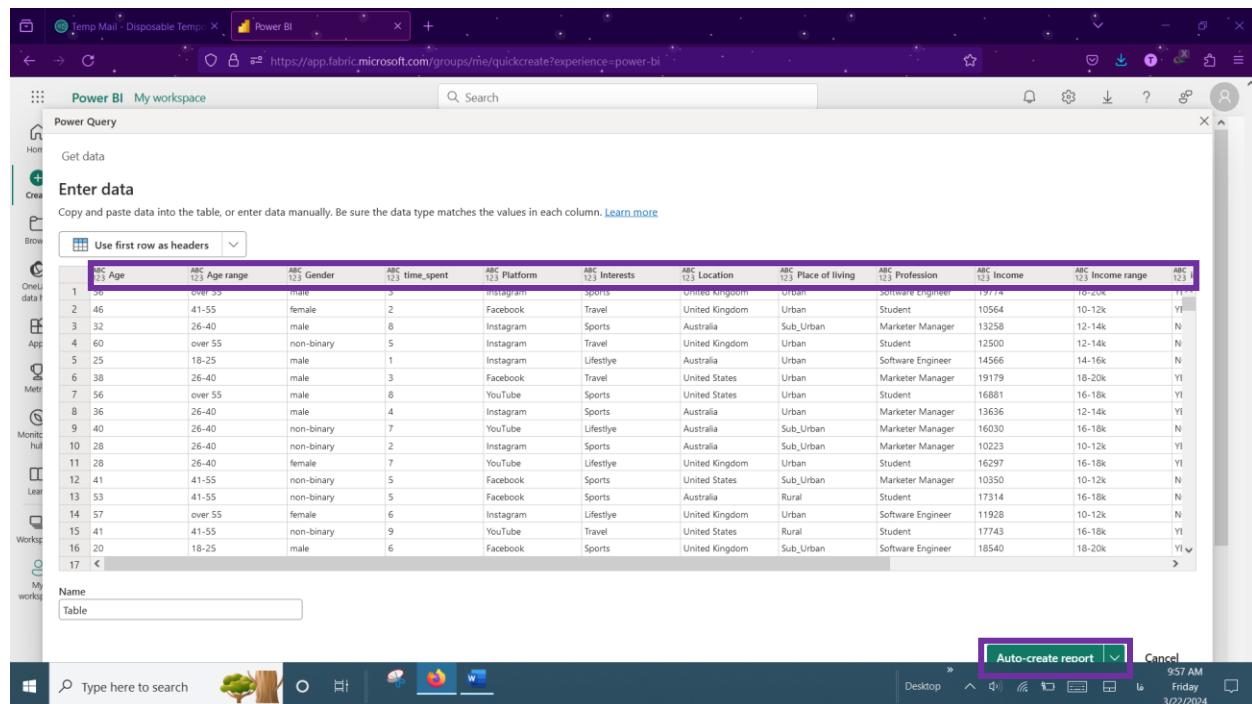
همهی Cell ها را انتخاب میکنیم به غیر از اسم ستون ها و کپی میکنیم.(تعداد دیتاهای ما زیاد هست و خوب شاید به همین دلیل است که در ابتدا انجام نمی شد.)

	D	E	F	G	H	I	J	K	L	M
1	time_spent	Platform	Interests	Location	Place of living	Profession	Income	Income range	indebt	isHomeOwner
2	3	Instagram	Sports	United Kingdom	Urban	Software Engineer	19774	18-20k	YES	NO
3	2	Facebook	Travel	United Kingdom	Urban	Student	10564	10-12k	YES	YES
4	8	Instagram	Sports	Australia	Sub_Urban	Marketer Manager	13258	12-14k	NO	NO
5	5	Instagram	Travel	United Kingdom	Urban	Student	12500	12-14k	NO	YES
6	1	Instagram	Lifestyle	Australia	Urban	Software Engineer	14566	14-16k	NO	YES
7	3	Facebook	Travel	United States	Urban	Marketer Manager	19179	18-20k	YES	YES
8	8	YouTube	Sports	United States	Urban	Student	16881	16-18k	YES	YES
9	4	Instagram	Sports	Australia	Urban	Marketer Manager	13636	12-14k	YES	NO
10	7	YouTube	Lifestyle	Australia	Sub_Urban	Marketer Manager	16030	16-18k	NO	NO
11	2	Instagram	Sports	Australia	Sub_Urban	Marketer Manager	10223	10-12k	YES	NO
12	7	YouTube	Lifestyle	United Kingdom	Urban	Student	16297	16-18k	YES	YES
13	5	Facebook	Sports	United States	Sub_Urban	Marketer Manager	10350	10-12k	NO	NO
14	5	Facebook	Sports	Australia	Rural	Student	17314	16-18k	NO	YES
15	6	Instagram	Lifestyle	United Kingdom	Urban	Software Engineer	11928	10-12k	NO	YES
16	9	YouTube	Travel	United States	Rural	Student	17743	16-18k	YES	YES
17	6	Facebook	Sports	United Kingdom	Sub_Urban	Software Engineer	18540	18-20k	YES	YES
18	8	Instagram	Travel	United States	Sub_Urban	Student	14732	14-16k	YES	YES
19	7	Instagram	Sports	United Kingdom	Rural	Marketer Manager	11017	10-12k	NO	YES
20	9	Facebook	Sports	United States	Sub_Urban	Student	12226	12-14k	NO	NO
21	4	Facebook	Sports	United States	Urban	Marketer Manager	17206	16-18k	NO	NO
22	2	Instagram	Sports	United Kingdom	Rural	Software Engineer	17170	16-18k	NO	NO
23	7	YouTube	Travel	Australia	Urban	Student	16207	16-18k	YES	NO
24	6	YouTube	Travel	United States	Rural	Marketer Manager	12419	12-14k	YES	YES
25	1	YouTube	Sports	Australia	Sub_Urban	Software Engineer	16240	16-18k	YES	YES
26	9	Instagram	Sports	United Kingdom	Rural	Software Engineer	11893	10-12k	YES	NO
27	9	Facebook	Travel	Australia	Rural	Student	15322	14-16k	YES	NO
28	2	Instagram	Lifestyle	Australia	Urban	Software Engineer	17718	16-18k	YES	NO
29	3	Facebook	Lifestyle	Australia	Sub_Urban	Marketer Manager	16878	16-18k	YES	NO
30	5	YouTube	Travel	Australia	Urban	Software Engineer	11125	10-12k	YES	NO
31	5	YouTube	Sports	Australia	Sub_Urban	Software Engineer	13117	12-14k	NO	NO
32	1	Facebook	Sports	United States	Urban	Marketer Manager	14161	14-16k	NO	NO
33	8	Facebook	Lifestyle	United States	Rural	Software Engineer	14042	14-16k	YES	NO

در قسمت دادن دیتا در Power Bi Microsoft Fabric باید Paste کنیم.

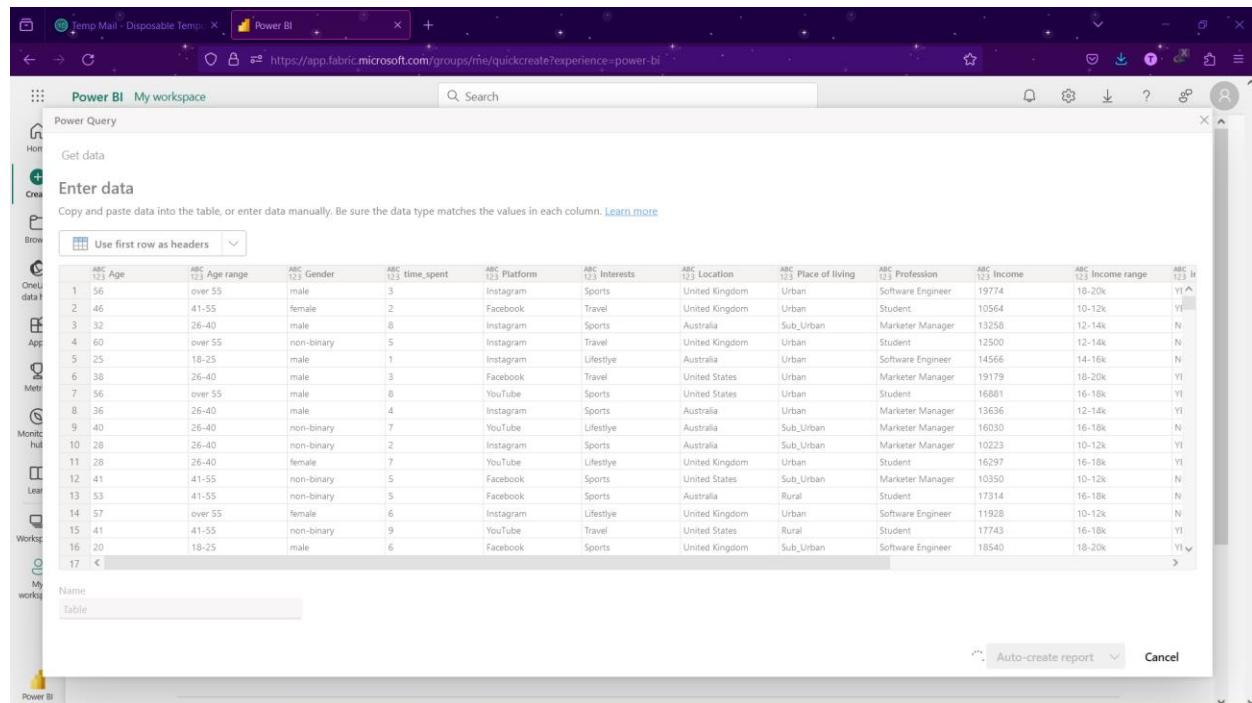
مشاهده میکنیم که اسم ستون ها نیست پس اسم ستون ها رو هم تک تک کپی و پیسٹ میکنیم در اینجا. روی Auto-

روی create report میز نیم.

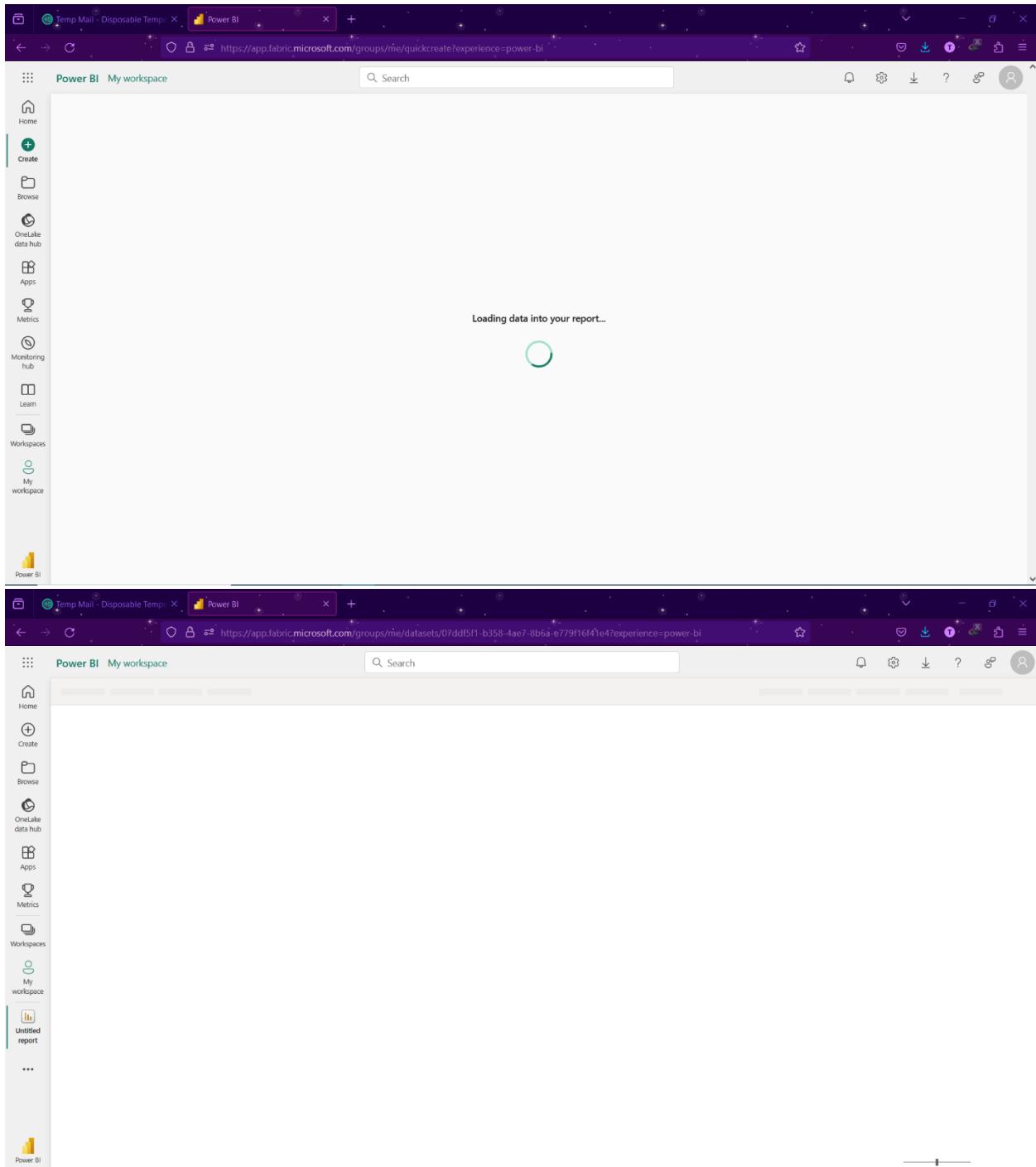


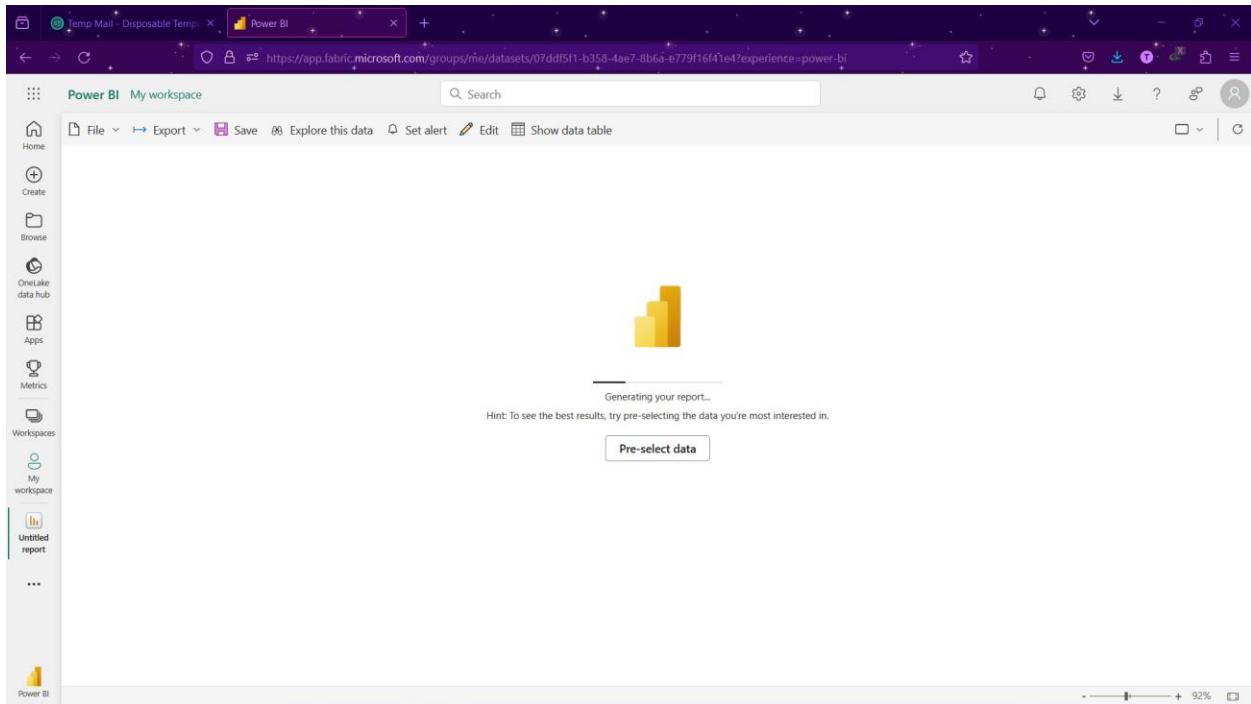
The screenshot shows the Power BI desktop interface with the 'Power BI My workspace' tab selected. In the center, there's a 'Enter data' section where a table of sample data is displayed. The table has columns: ABC 123 Age, ABC 123 Age range, ABC 123 Gender, ABC 123 time_spent, ABC 123 Platform, ABC 123 Interests, ABC 123 Location, ABC 123 Place of living, ABC 123 Profession, ABC 123 Income, ABC 123 Income range, and ABC 123 Ir. The data consists of 17 rows of user information. At the bottom right of the data entry area, there is a button labeled 'Auto-create report' with a dropdown arrow, which is highlighted with a yellow box.

صبر میکنیم.

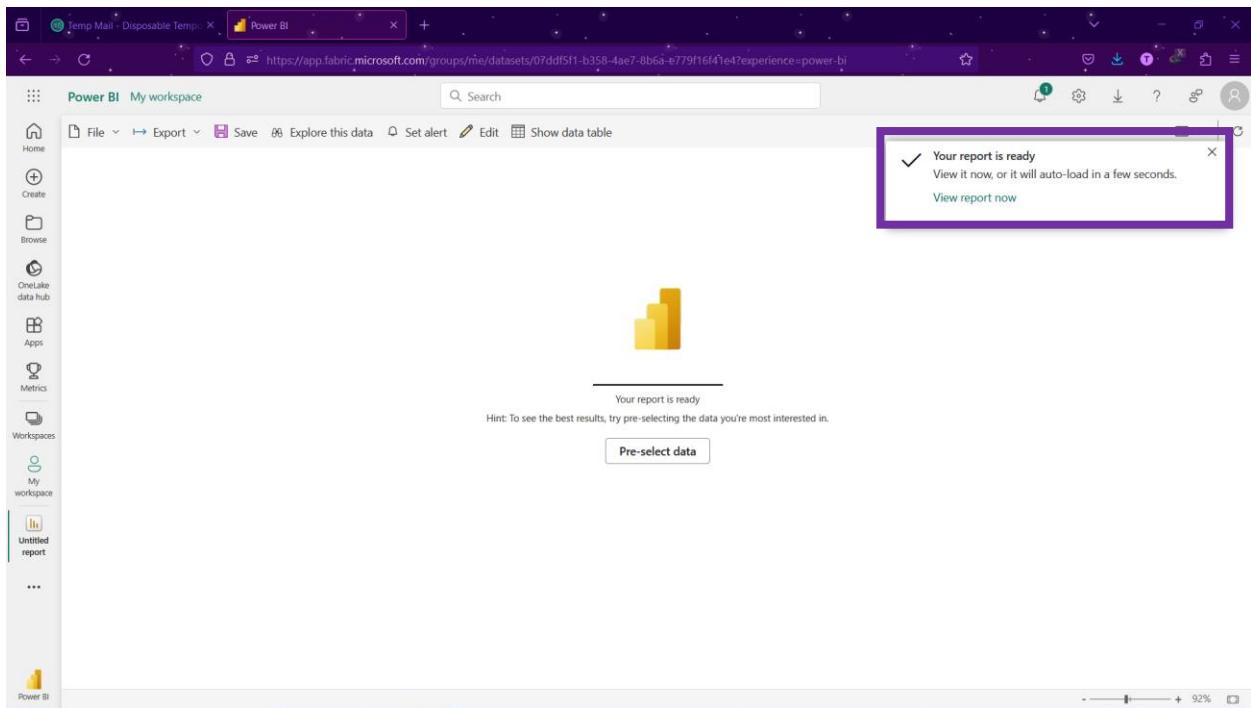


This screenshot is identical to the one above, showing the Power BI desktop interface with the 'Power BI My workspace' tab selected. It displays the 'Enter data' screen with the same table of sample data. The 'Auto-create report' button at the bottom right is again highlighted with a yellow box.

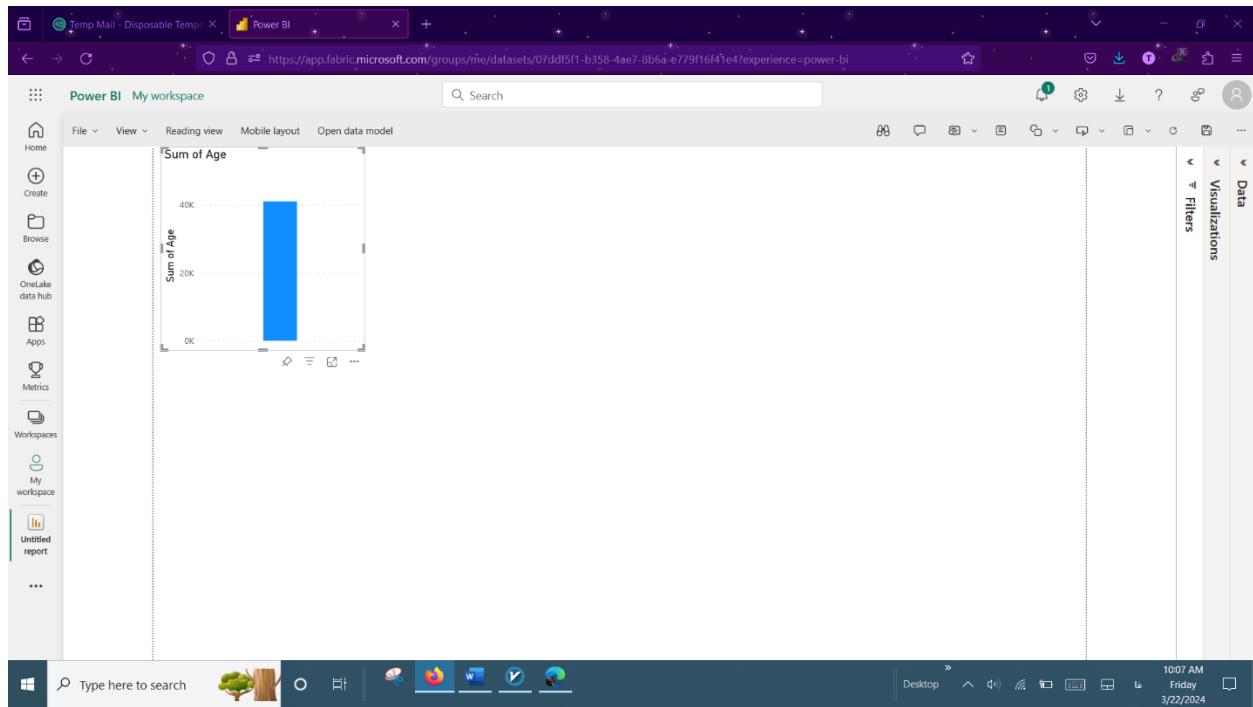




خب در اینجا report ما آماده شد.

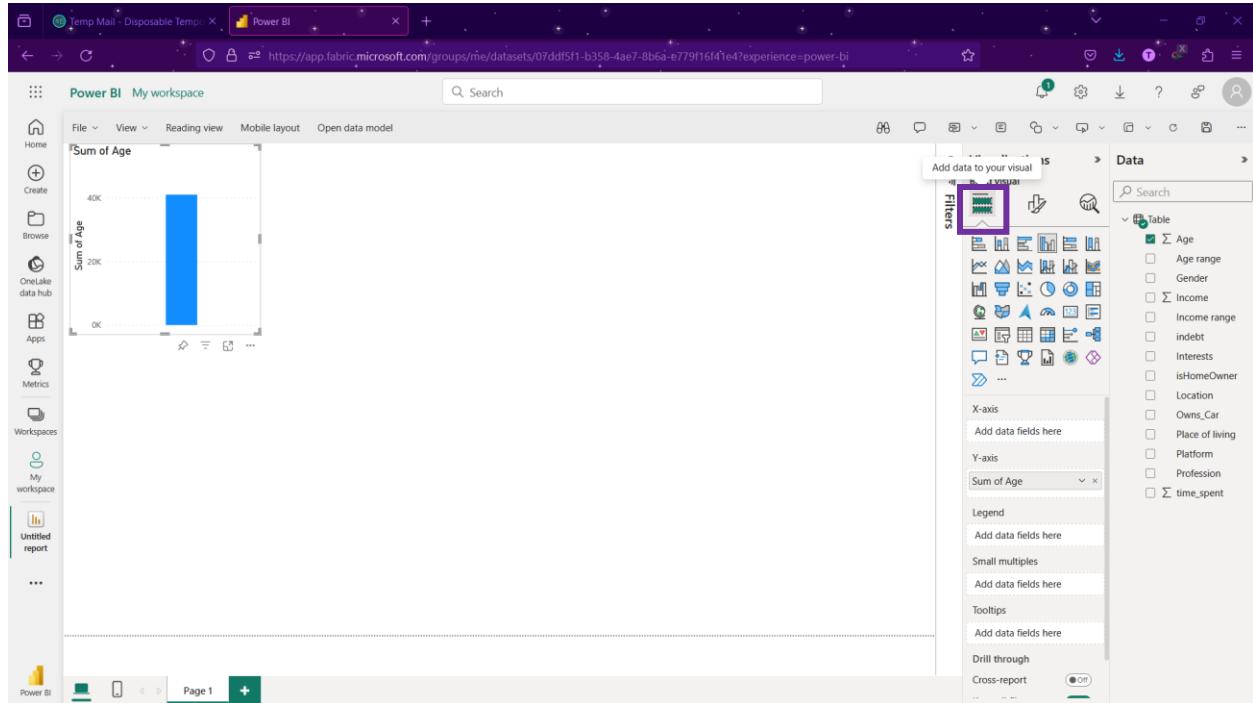


مانند Data, Visualization, Filters Microsoft Power BI در اینجا هم داریم.

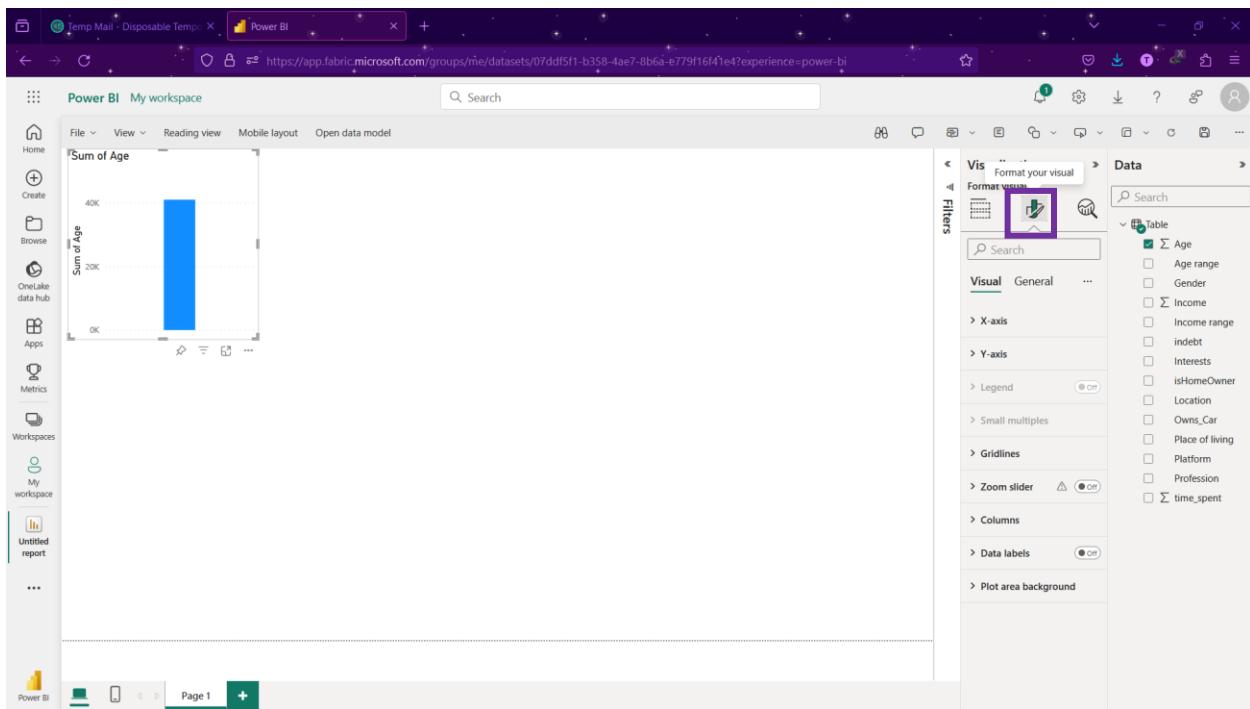


هایی از جدول ها قرار بگیرند که در نمودارها رسم شوند.

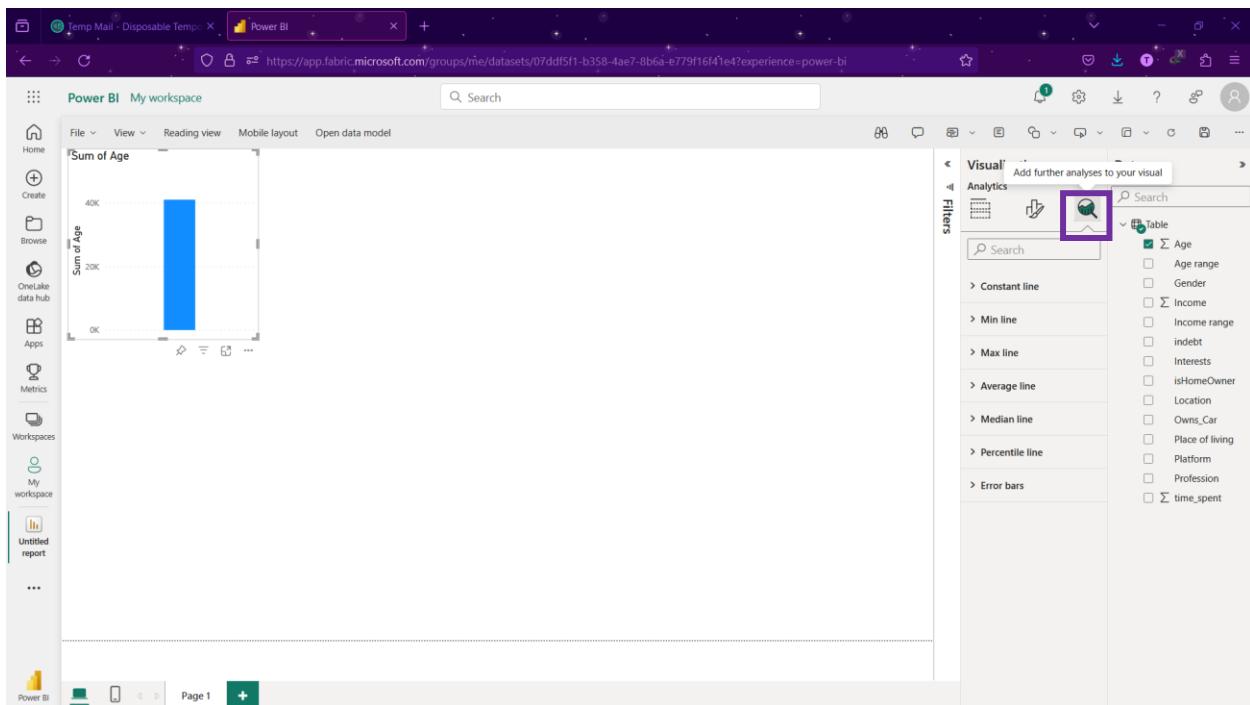
تو Fileds انتخاب میکنیم که چه Fileds هایی از جدول ها قرار بگیرند که در نمودارها رسم شوند.



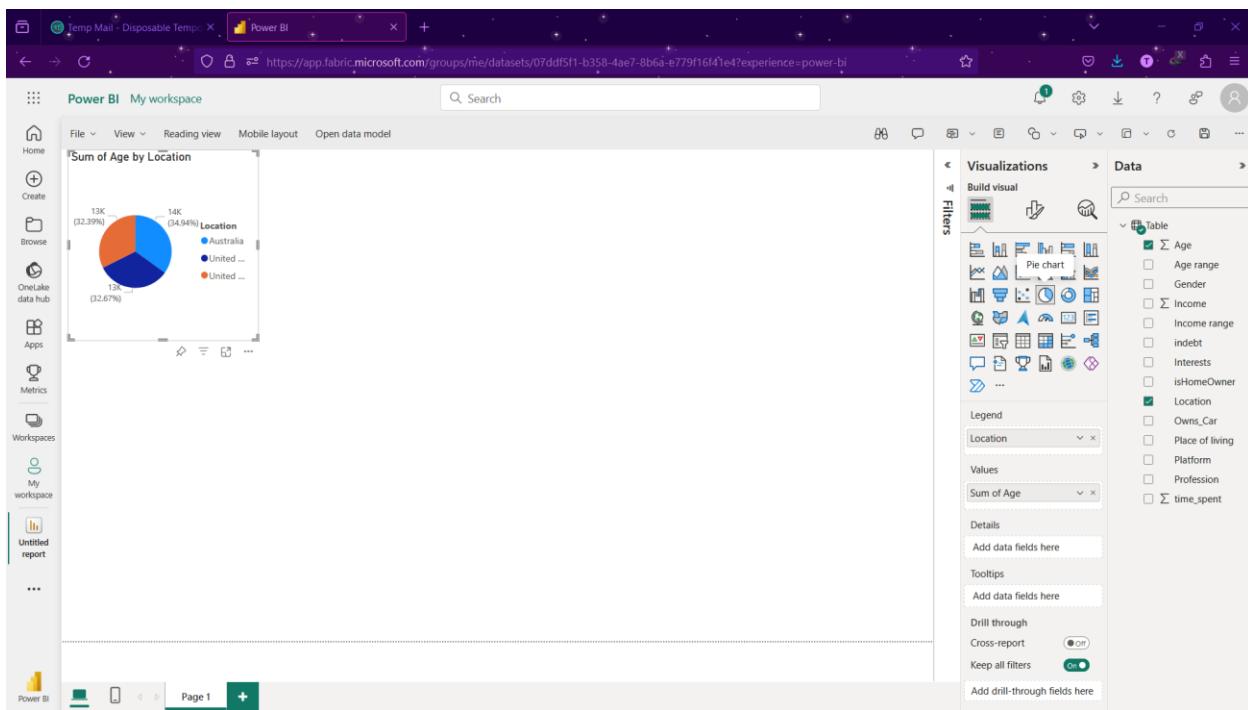
تو Format Format .2: انتخاب میکنیم چه رنگی و چه فونتی و چه سایزی باشه(تنظیمات ظاهری)



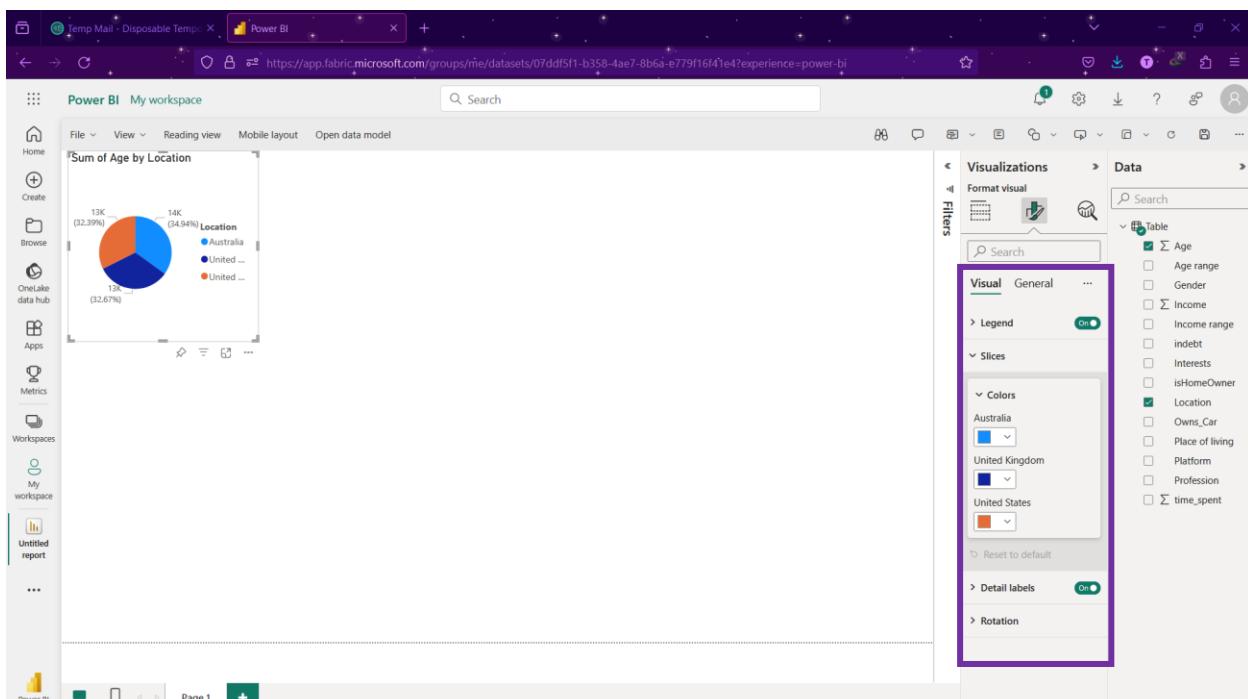
یک سری خط و خطوط مثل trend line , average line , forecast , ... رسم میکنیم. Analyses .3



نمودار دایره ای رو انتخاب میکنیم و سن و مکان رو انتخاب میکنیم. که به صورت زیر به ما نمایش می دهد.

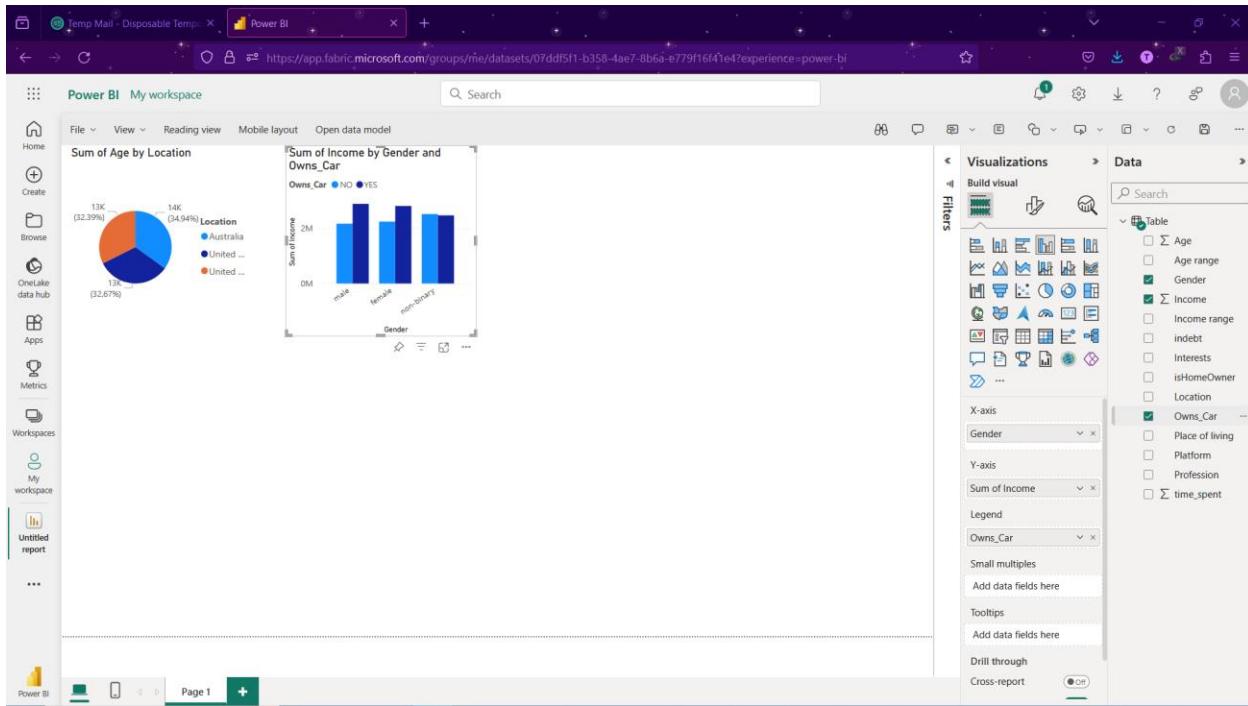


برای تنظیمات ظاهری میتوان به قسمت فرمت رفت و رنگ و سایز و ... رو تغییر داد.

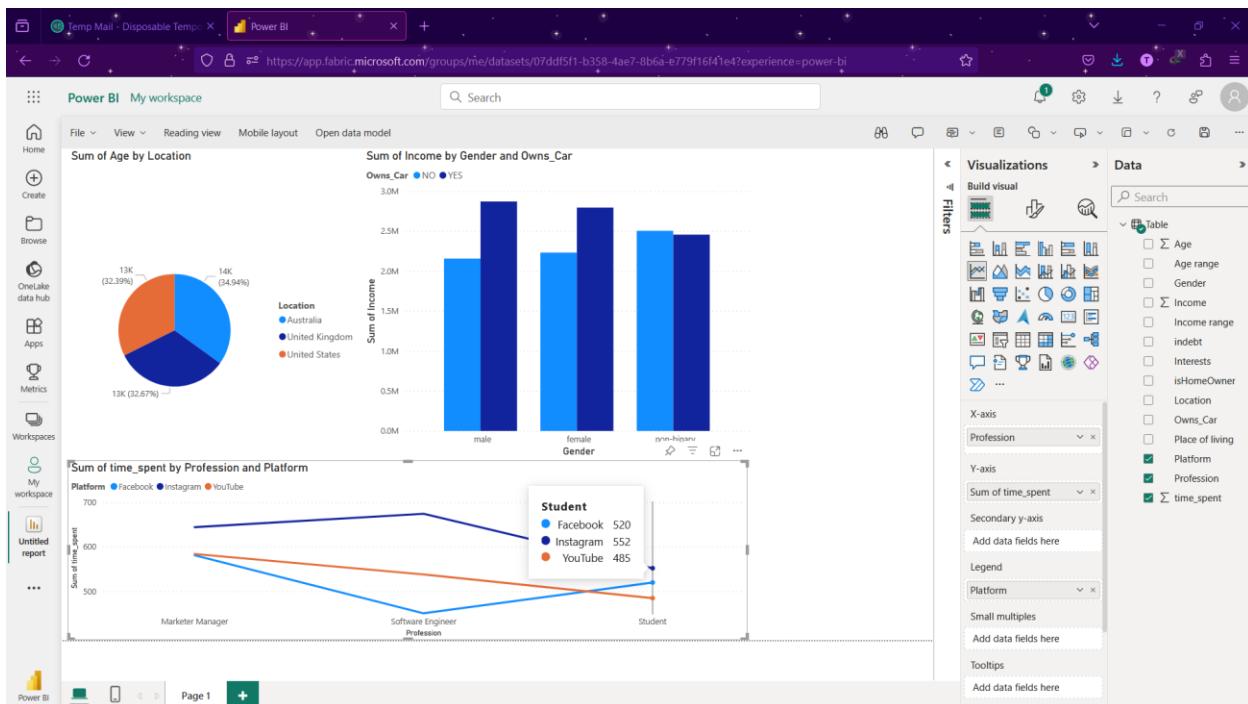


روی **clustered column chart** کلیک میکنیم.(نمودار مبله ای)

درآمد رو بر حسب جنسیت و اینکه ماشین از خودشون دارن یا نه مورد بررسی قرار میدهیم.

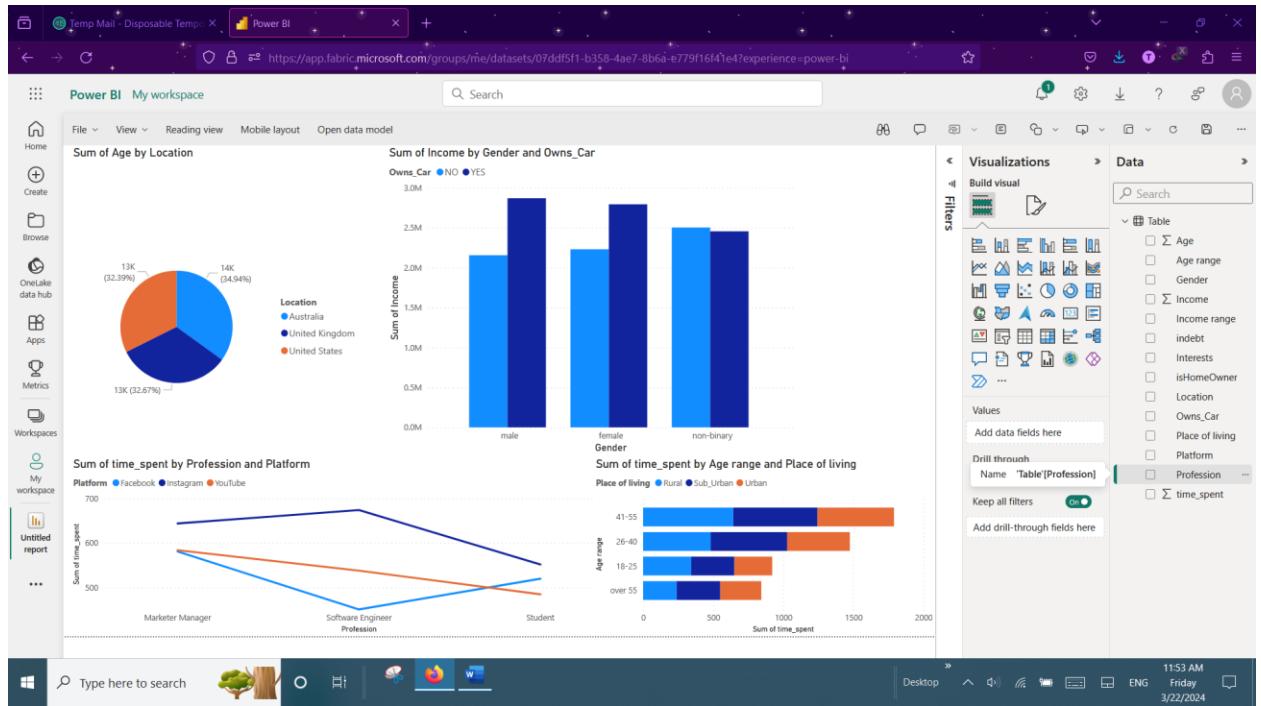


نمودار خطی رو هم انتخاب میکنیم و زمان سپری شده بر حسب تخصص یا حرفه و پنفرمی که زمانشون رو سپری میکنند انتخاب می کنیم.



روی هر قسمت هم نگه داریم به صورت دقیق بهمون عدد و اسم اون پلتفرم رو میده.

نمودار میله ای انباشه رو هم برای بررسی زمان سپری شده بر حسب سن و محل زندگی ایجاد میکنیم.



مقایسه Microsoft Fabric & Microsoft Power BI

خب بعد اینکه پروژه ها را در دو ابزار پیاده سازی کردیم به مقایسه آنها می پردازیم. براساس تجربه یا براساس دانشی که توضیحات جفتشون داشتن کلیت قابلیت جفتشون رو مقایسه کردیم.

از نظر کامل قابلیت ها مقایسه شود خیلی بهتره چون جامعیت بیشتری دارد.

Microsoft Fabric

مزایا:

- در Microsoft Fabric ما به صورت اتوماتیک دیتا رو که دادیم خودش برآمون داشبورد درست کرد و خب به راحتی به ساختن نمودارها و انتخاب دیتاهای مورد نظر می پردازیم.
- آنلاین هست. امکاناتی از قبیل Share با افراد دیگر و ... را در اختیار ما می گذارد.
- فابریک قدرت هاش خیلی بیشتره واقعاً نسبت به پاور بی آی. بخش trailer free ش رو داریم استفاده میکنیم.
- یه بخش هایی از پاور بی آی رایگان هم هست ولی فابریک کلا باید اشتراک داشته باشیم.

معایب:

- ولی خب در ابتدا به یک ایمیل کاری نیاز هست که ما چون دانشگاه بهمون ایمیل کاری نمی داد یکم اذیت شدیم و بعد با استفاده از temp mail به ادامه کار در فابریک پرداختیم.
- در حالیکه در Microsoft Fabric ما به دلیل اینکه فقط با ایمیل کاری میشه وارد وبسایت شد و خب احتیاج به اطلاعات شرکت و تایید آن داشت لاینسن معتری نداشتم، کار زیادی نمیتوانستیم بکنیم و خب خیلی محدود بودیم. حتی برخی از نمودارها هم در دسترس ما نبود.
- تعداد دیتاهایی که میتوانستیم بدیم خیلی محدود بود. اگر دیتاست دیتاهاش تعدادشون زیاد باشه قبول نمیکنه و خب یک سری از دیتاهای دور ریخته می شوند.

Microsoft Power BI

مزایا:

- Microsoft Power BI احتیاجی به ایمیل کاری نداشت و امکانات بیشتری در دسترس ما بود.
- در هر لحظه که سیستم را روشن کنیم بدون اتصال به اینترنت در دسترس ماست.
- برای کارهای معمولی کار ما را راه می اندازد.

معایب:

- باید نرم افزار دانلود شود و حجمی از سیستم را اشغال می کند.
- به صورت آفلاین استفاده میکنیم حالا قابلیت این رو نداره که بخواهیم با ایمیل شیر کنیم. جایی بخوایم با تیم یا سازمانمن شیر کنیم این قابلیت ها رو به ما نمیده.
- برای اینکه حرfe ای کار کنیم احتیاج به اشتراک هست.

در تجربه ای که کسب کردیم:

- فابریک فوق العاده دنیای بزرگی دارد. هرچند که ما در ایران هستیم و محدودیت داریم. محدودیت بیشتری داریم.
 - مزیتش این نیست که تنها یک داشبورد اتومات برای ما درست کند، کلا برای کار با دیتابست های بزرگ درست شده.
 - یک حالت یکپارچه شدست برای کسایی که مثلا میخوان دیتا آنالیست کار کنن، کارای علم داده انجام بدن و
 - پاور بی آی فقط یک جز کوچیک از مجموعه ابزارهای فابریکه که خیلی راحت یک حالت n تو n داره.(سرچ کنیم فابریک. تنوی بخش تصاویر برمی، چندتا مربع هست هرکدام یکی از بخش های فابریکه از یکی فلاش خورده به اون یکی. یعنی تا این حد پویا میکنه پروژه رو که میتوانیم انبار داده های خیلی بزرگ داشته باشیم که داخل onelake میداره.)
 - Clean data, create model کرده.
 - اجرаш هم برای علم داده و مدلسازی و ... برد میخوره. یک جورایی ماشین لرنینگ هست.
 - همه اینها رو کنار همیگه جمع کرده و در نهایت میزنه به پاور بی آی که بتونه گزارش گیری انجام بشه و این گزارش گیری به جای اینکه تنها گزارش گیری باشد یک حالت پیش بینی طوری میکند.
 - جفشنون پولیه برای کار حرفه ای
 - آفلاین و آنلاین میشود.
 - ولی فابریک خیلی بالاتر قدرتش
- در نتیجه با بررسی فعالیت و خواسته هایی که هست میتوان انتخاب کرد که کدام یک برای ما بهتر است.
- داشبورد داریم می سازیم با توجه به پروژه ای که داریم.
- قیاسش اینه که پاور بی آی بخشی از فابریکه که اون اخر به ما نشون میده که چه اتفاقی میوفته.

منابع

- ❖ Skool.com/Microsoft-fabric
- ❖ Microsoft Power BI
- ❖ Microsoft Fabric

"پاشکر از استاد عزیزمون سرکار خانم دکتر ملک زاده  با آرزوی سلامتی،
....."

پیشرفت و بهترین ها در سال جدید"