ساختار پروژه

1. توسعه و پیاده‌سازی ساختار اولیه گزارش – 1.1 آشنایی با داده‌ها بررسی اولیه داده‌ها و شناخت کلی از محتوا و ساختار آن‌ها - 1.2 ایجاد زیرساخت پروژه ساخت تمپلیت اولیه، تعریف ساختار پوشه‌ها و راه‌اندازی مخزن گیت - 1.3 تحلیل آماری اولیه داده‌ها انجام تحلیل‌های پایه شامل: تعداد ردیف‌ها و ستون‌ها تعداد و درصد داده‌های خالی بررسی کلی توزیع داده‌ها - 1.4 بررسی نوع داده‌ها و کاربرد آن‌ها مشخص کردن نوع هر فیلد (عددی، متنی، تاریخ و ...) و توضیح نقش و کاربرد هر کدام در تحلیل - 1.5 شناسایی و تعریف KPI ها مشخص کردن شاخص‌های کلیدی عملکرد و توضیح اهمیت هر کی پی ای - 1.6 تعیین هدف تحلیل داده تعریف هدف اصلی تحلیل و تصمیم‌هایی که قرار است بر اساس نتایج آن گرفته شود - 1.7 طراحی رویکرد بصری‌سازی داده بررسی روش‌های مناسب برای نمایش داده‌ها به‌صورت بصری، ساده و قابل فهم - 1.8 تعریف سوالات تجاری و محدودیت‌ها شناسایی سوالات اصلی کسب‌وکار بررسی ریسک‌ها و محدودیت‌های داده تعیین نوع گزارش مناسب برای مخاطبان مختلف - 1.9 مستندسازی و ثبت در Git ثبت نتایج این مرحله، تهیه گزارش و انجام Commit در مخزن Git.

2. مهندسی داده و مدل‌سازی – ضبط ویدیو از طول فرایند اجرایی - 2.1 انجام ETL با Power Query پاک‌سازی، تبدیل و آماده‌سازی داده‌ها برای استفاده در گزارش - 2.2 مستندسازی فرآیند ETL ثبت مراحل انجام‌شده در ETL و توضیح تغییرات اعمال‌شده روی داده‌ها - 2.3 طراحی مدل داده طراحی ساختار مناسب داده‌ها با رویکرد Star Schema برای بهبود خوانایی و عملکرد - 2.4 تعریف ارتباط بین جداول ایجاد و تنظیم Relationships بین جداول به‌صورت صحیح و بهینه - 2.5 ایجاد Date Table و محاسبات پایه ساخت جدول تاریخ استاندارد و تعریف Measures اولیه مورد نیاز تحلیل - 2.6 ثبت تغییرات در Git انجام Commit و ثبت وضعیت پروژه پس از تکمیل این مرحله.

3. طراحی تجربه کاربری و داشبورد - ضبط ویدیو از طول فرایند اجرایی - 3.1 شناسایی سوالات اصلی استخراج مهم‌ترین سوال‌هایی که گزارش باید به آن‌ها پاسخ دهد - 3.2 دسته‌بندی سوالات گروه‌بندی سوالات بر اساس موضوع و اولویت برای طراحی بهتر گزارش - 3.3 طراحی داستان داده تعریف سناریوی داستان‌پردازی برای نمایش داده‌ها به‌صورت مرحله‌به‌مرحله و قابل فهم - 3.4 انتخاب و آماده‌سازی قالب گزارش انتخاب تمپلیت مناسب و هماهنگ‌سازی آن با رنگ‌ها و هویت برند -3.5 طراحی ساختار کلی داشبورد مشخص کردن چیدمان کلی داشبورد و محل قرارگیری اجزای اصلی - 3.6 طراحی صفحات و نمودارها و کمک از تمپلیت خودم طراحی هر صفحه و انتخاب نوع نمودار متناسب با داده‌ها و سوالات - 3.7 ثبت تغییرات در Git انجام Commit پس از تکمیل این مرحله و ثبت وضعیت پروژه.

4. (Implementation)پیاده‌سازی گزارش - 4.1 ضبط ویدیو از فرآیند اجرا ضبط ویدیو از مراحل اجرایی پروژه برای مستندسازی و ارائه روند کار - 4.2 پیاده‌سازی اجزای مشترک صفحات ایجاد و پیاده‌سازی اجزای یکسان در صفحات مختلف مانند:گروه‌بندی المان‌ها، بوکمارک ها، ساختارهای تکرار شونده - 4.3 پیاده‌سازی صفحات و نمودارها پیاده‌سازی هر بخش از صفحات و رسم نمودارها؛ پس از تکمیل هر صفحه، ثبت Commit انجام می‌شود - 4.4 بازنگری فنی اولیه بررسی و اصلاح موارد فنی شامل: رنگ‌ها فونت‌ها اندازه و چیدمان المان‌ها - 4.5 بازنگری داستان‌پردازی و نمودارها بازبینی نمودارها و بهبود جریان داستان داده برای انتقال بهتر پیام - 4.6 به‌روزرسانی نهایی و ثبت در Git اعمال اصلاحات نهایی و انجام Commit برای ثبت نسخه پایدار گزارش.

5. معرفی و ارائه پروژه - 5.1 تهیه اسلاید معرفی پروژه آماده‌سازی یک اسلاید متنی برای معرفی کلی پروژه، هدف آن و خروجی‌ها - 5.2 تولید محتوای معرفی پروژه تهیه یک پست شامل: توضیح فرایند انجام پروژه ویرایش محتوای بصری نوشتن متن توضیحی (Caption) برای انتشار - 5.3 ضبط ویدیو معرفی به زبان انگلیسی ضبط یک ویدیو به زبان انگلیسی شامل ارائه پروژه، توضیح رویکرد کاری و دلایل انتخاب هر بخش.

1. توسعه و پیاده‌سازی ساختار اولیه گزارش

1.1.1 تشخیص نقش جدول Fact و Dimension ؟

جدول Fact چیست؟ Fact Table جدولِ «اتفاق‌ها / تراکنش‌ها» است ویژگی‌های کلیدی تعداد رکورد زیاد، شامل اعداد قابل محاسبه، هر ردیف = یک رویداد (مثلاً یک فروش)، معمولاً شامل کلید خارجی به Dimension ها و جدول Dimension چیست؟ Dimension Table جدولِ «توصیف‌کننده» است ویژگی‌های کلیدی تعداد رکورد کم‌تر، اطلاعات متنی و توصیفی، برای فیلتر، گروه‌بندی و Slice کردن داده‌ها، هر Dimension معمولاً یک Primary Key دارد با بررسی داده‌های ارائه‌شده در این پروژه، می‌توان نتیجه گرفت که جدول «فروش» نقش جدول Fact و جدول «مشترکین» نقش جدول Dimension را ایفا می‌کند. دلیل این دسته‌بندی آن است که کلیه داده‌های اصلی و تراکنشی سیستم در جدول فروش قرار گرفته‌اند و هر ردیف از این جدول نشان‌دهنده یک رویداد مشخص فروش یا صدور فاکتور است. همچنین این جدول شامل مقادیر عددی قابل محاسبه مانند مبلغ، هزینه، تخفیف و مالیات بوده و مبنای اصلی تعریف شاخص‌های کلیدی عملکرد و محاسبات تحلیلی در داشبورد Power BI محسوب می‌شود. در مقابل، جدول مشترکین شامل اطلاعات توصیفی مربوط به مشتریان است که به‌منظور شناسایی، دسته‌بندی و فیلتر کردن داده‌های فروش مورد استفاده قرار می‌گیرد. این جدول دارای داده‌هایی مانند مشخصات فردی، موقعیت جغرافیایی و نوع مشتری بوده و فاقد مقادیر عددی تراکنشی است، بنابراین به‌عنوان جدول Dimension شناخته می‌شود. ارتباط بین این دو جدول از طریق کلید مشترک «کد مشترک» برقرار شده و مدل داده به‌صورت یک مدل ستاره‌ای با جدول Fact در مرکز و جدول Dimension در اطراف آن طراحی می‌شود.

1.1.2 داده ها در چه نوع صنعت است و کمی کسب اطلاع از ساختار و صحبت با مجموعه ؟

با بررسی ماهیت داده‌های موجود در این پروژه می‌توان نتیجه گرفت که این داده‌ها متعلق به صنعت خدمات و به‌طور مشخص حوزه مخابرات و ارائه‌دهندگان خدمات ارتباطی است. وجود مفاهیمی مانند مشترک، نوع سرویس، نوع خدمات، پکیج، ARPU، دوره‌های صورتحساب، تاریخ شروع و پایان فاکتور و همچنین فاکتورهای فروش دوره‌ای نشان می‌دهد که ساختار داده‌ها مبتنی بر ارائه خدمات اشتراکی بوده و درآمدزایی به‌صورت دوره‌ای انجام می‌شود. در چنین صنایعی تمرکز اصلی بر مدیریت مشترکین، تحلیل رفتار مصرف، بررسی میزان درآمد هر مشترک و ارزیابی اثربخشی پکیج‌ها و سرویس‌ها است. برای کسب شناخت بهتر از ساختار داده‌ها، لازم است اطلاعات اولیه‌ای از فرآیندهای کسب‌وکار مجموعه جمع‌آوری شود که این امر معمولاً از طریق بررسی مستندات داخلی و گفتگو با ذی‌نفعان مجموعه مانند واحد فروش، مالی و پشتیبانی انجام می‌شود. در این گفتگوها مواردی مانند نحوه صدور فاکتور، تعریف دوره‌های مالی، ساختار قیمت‌گذاری خدمات، دسته‌بندی مشترکین و شاخص‌های کلیدی عملکرد مورد توجه قرار می‌گیرد تا داشبورد طراحی‌شده با نیازهای واقعی سازمان هم‌راستا بوده و تحلیل‌ها از اعتبار و کاربردپذیری بالاتری برخوردار باشند.

1.1.3 بررسی صد الی دویست دریف از ستون ها برای درک بهتر داده ها و مشاهده نقض و مشکل در داده ها ؟

با بررسی نمونه‌ای از داده‌ها در حدود صد تا دویست ردیف از جدول فروش و مشترکین، ساختار کلی داده‌ها و کیفیت ثبت اطلاعات مورد ارزیابی قرار گرفت. در جدول فروش، هرچند داده‌های تراکنشی به‌صورت کامل‌تری نسبت به سایر جداول ثبت شده‌اند، اما برخی نواقص قابل مشاهده است. از جمله این نواقص می‌توان به وجود مقادیر خالی در برخی ستون‌های توصیفی مانند نوع سرویس، نوع خدمات یا پکیج اشاره کرد که می‌تواند در تحلیل‌های مبتنی بر دسته‌بندی اختلال ایجاد کند. همچنین در برخی رکوردها ناسازگاری در فرمت تاریخ‌ها مشاهده می‌شود، به‌طوری که تاریخ شروع و پایان فاکتور در برخی موارد هم‌خوانی منطقی ندارند یا به‌صورت یکنواخت ثبت نشده‌اند. علاوه بر این، وجود مقادیر عددی صفر یا غیرمنطقی در برخی ستون‌های مالی مانند مبلغ یا هزینه می‌تواند نشان‌دهنده ثبت ناقص فاکتور یا خطای سیستمی در فرآیند صدور باشد. (در صورت مشاهده هرگونه نواقص با کارفرما در میان گذاشته و دلیل هر گونه مشکل را پیدا کرده)

در جدول مشترکین نیز بررسی داده‌ها نشان می‌دهد که برخی فیلدهای اطلاعاتی مانند جنسیت، نوع مشتری یا موقعیت جغرافیایی به‌صورت ناقص یا ناهماهنگ ثبت شده‌اند و در بعضی موارد یک مفهوم واحد با مقادیر متنی متفاوت نمایش داده شده است که این

موضوع نیازمند استانداردسازی است. همچنین مشاهده می‌شود که برخی مشترکین فاقد ارتباط کامل با جدول فروش هستند یا بالعکس، رکوردهایی در جدول فروش وجود دارد که اطلاعات متناظر آن‌ها در جدول مشترکین ناقص یا غیرفعال است.
در پایان این مرحله، با طرح سوالات تحلیلی اولیه از داده‌ها، مانند کامل بودن کد مشترک در تمامی فاکتورها، سازگاری بازه‌های زمانی فاکتور با نوع سرویس، و اعتبار مقادیر مالی ثبت‌شده، مجموعه‌ای از نواقص و مشکلات داده‌ای شناسایی شد. این موارد به‌عنوان ورودی مرحله پاک‌سازی و پیش‌پردازش داده‌ها در Power Query ثبت می‌شوند تا پیش از طراحی مدل نهایی و محاسبه شاخص‌های تحلیلی، کیفیت داده‌ها به سطح قابل قبول برای تحلیل و تصمیم‌گیری ارتقا یابد.

1.1.4 در اخر پرسیدن سوال از داده ها و یادداشت هرگونه نقض و مشکل در داده ها ؟

من در جلسه که با کارفرما داشتم تمام مشکلات را مطرح کروم در بعضی از داده ها مشتری برای تست اتصال می داد و این مقادیر به همین علت صفر شده ایا باید پاک شود خیر به این علت که کارفرما می خواهد بداند چه زمانی این مشکل بود برای عیب یابی با چند درصد مشتری ها این مشکل را داشته اند و برای ستون به نام تاریخ شروع فاکتور مقادیر بلانک هست به این علت در سیستم مشکل به وجود آمده و میخواهد با توجه به اطلاع دیکر رمان درخواست را ببیند و مشکل را عیب یابی کند و یک نکته مهم ما این داده ها را پاک نمی کنیم به این علت که یک ستون خالی اما در همان ردیف داده های مهم دیگری دارد در یک صورت پاک می کنیم که داده های مهم خالی یا نال باشد

1.2.1 استفاده از ساختار تعیین شده برای پروژه شامل - 1.2.2 بعد کپی از ساختار ئ تعیین اسم پروژه پروژه را در گیت هاب قرار داده و اولین کامیت را انجام داده

ساختار به دقت انجام شد و در گیت هاب قرار گرفت برای دیدن در لینک متن گیت هاب کلیک کرده

Data	1/31/2026 3:49 PM	File folder	
Docs	1/5/2026 5:04 PM	File folder	
Images	1/31/2026 3:47 PM	File folder	
Media	1/31/2026 3:49 PM	File folder	
Report	12/6/2025 11:40 AM	File folder	
Theme	1/31/2026 3:46 PM	File folder	
.gitignore	12/15/2025 3:41 PM	Git Ignore Source ...	1 KB
README.md	1/5/2026 5:06 PM	Markdown Source ...	0 KB

1.3.1 در این قسمت من باید یک تحلیل از داده ها به دست آورم (راهنما : با کمک پاور بی ای و دکس من میتونم به راحتی این تحلیل را انجام دهم) ؟

در این مرحله، با استفاده از قابلیت‌های تحلیلی Power BI و محاسبات آماری قابل استخراج از داده‌ها، یک تحلیل اولیه به‌منظور درک رفتار داده‌ها، شناسایی الگوها و کشف ناهنجاری‌های احتمالی انجام شد. بررسی جدول Fact_Sales نشان می‌دهد که این جدول شامل 38657 رکورد فروش بوده و هیچ مقدار Null در ستون‌های کلیدی آن مشاهده نمی‌شود که این موضوع نشان‌دهنده کامل بودن ثبت تراکنش‌ها از منظر ساختاری است. ستون مبلغ خالص دارای حداقل مقدار صفر و حداکثر مقدار 75,400,000 است که دامنه گسترده‌ای از مبالغ فروش را نشان می‌دهد و بیانگر تنوع بالای فاکتورها از نظر ارزش مالی است. مقدار میانگین مبلغ خالص حدود 1,381,186 و انحراف معیار نسبتاً بالا است که حاکی از پراکندگی زیاد داده‌ها و وجود فاکتورهایی با مقادیر بسیار بزرگ‌تر از میانگین می‌باشد. وجود تعداد قابل توجهی مقدار صفر در ستون مبلغ خالص و هزینه می‌تواند نشان‌دهنده فاکتورهای اصلاحی، تستی یا ثبت ناقص برخی تراکنش‌ها باشد که نیازمند بررسی بیشتر است.

در ستون تخفیف، وجود مقادیر منفی تا حداقل 12,993,500- نشان‌دهنده اعمال اصلاحات مالی، برگشت از فروش یا مکانیزم‌های تعدیل حساب است که در تحلیل‌های مالی باید به‌صورت جداگانه در نظر گرفته شوند. همچنین میانگین نزدیک به صفر در این ستون بیانگر آن است که بخش عمده‌ای از فاکتورها بدون تخفیف ثبت شده‌اند. ستون مالیات نیز دارای مقادیر منفی و مثبت است که می‌تواند ناشی از اصلاح فاکتورها یا برگشت مالیات باشد و نیازمند تعریف منطق مشخص در محاسبات DAX است. بررسی ستون کد مشترک در جدول فروش نشان می‌دهد که 10310 مشترک یکتا در این جدول وجود دارد که این موضوع بیانگر آن است که هر مشترک به‌طور میانگین بیش از یک فاکتور در بازه زمانی مورد بررسی داشته است.

از منظر زمانی، داده‌های جدول فروش بازه‌ای از تاریخ 1402/11/01 تا 1403/03/15 را پوشش می‌دهند که امکان تحلیل روندهای کوتاه‌مدت فروش، تهدیدها و رفتار مشترکین در این بازه را فراهم می‌کند. ستون‌های توصیفی مانند نوع استفاده، نوع سرویس، نوع خدمات، پکیج و وضعیت دارای تعداد مقادیر یکتای محدود هستند که این ویژگی آن‌ها را برای استفاده در فیلترها، Slicer ها و تحلیل‌های مقایسه‌ای بسیار مناسب می‌سازد.

در جدول Dim Subscribers، تعداد 40820 مشترک یکتا ثبت شده است که بیشتر از تعداد مشترکین حاضر در جدول فروش است و این موضوع نشان می‌دهد که بخشی از مشترکین در بازه زمانی تحلیل، فاکتور فعالی نداشته‌اند. ستون ARPU در این جدول دارای دامنه بسیار گسترده‌ای از مقادیر منفی تا مثبت بوده و انحراف معیار بالایی دارد که می‌تواند نشان‌دهنده ناهنجاری در محاسبه ARPU، وجود مشترکین غیرفعال یا ثبت داده‌های غیرعادی باشد. همچنین وجود تاریخ‌های نامعتبر یا بسیار قدیمی در ستون تاریخ تولد، مانند سال 1841، نشان‌دهنده داده‌های پیش‌فرض یا خطای ورود اطلاعات است که باید در مرحله پاک‌سازی داده‌ها اصلاح یا فیلتر شوند.

در مجموع، این تحلیل اولیه نشان می‌دهد که داده‌ها از نظر ساختاری کامل هستند اما از نظر محتوایی دارای پراکندگی بالا، مقادیر غیرمنطقی و ناهنجاری‌هایی هستند که شناسایی آن‌ها در این مرحله، زمینه را برای تعریف Measure های دقیق در DAX، پاک‌سازی داده‌ها و طراحی داشبورد تحلیلی معتبر و قابل اتکا فراهم می‌کند.

- 1.3.2 یک توزیع از داده های عددی هم بدست بیارم به کمک داده ها و نمودار های خود پاور بی ای

1.4.1 در این مرحله تو تحلیل نمی‌کنی، فقط: داده را می‌شناسی نقش هر ستون را مشخص می‌کنی تصمیم می‌گیری با هر ستون چه کار کنی و خروجی را برای هر ستون در جدول زیر پر کن

Column Name	Data Type	Role	Usage
"مبلغ خالص"	Whole Number	Measure (Fact Attribute)	استفاده برای محاسبه مجموع فروش، میانگین فروش، KPIهای درآمدی
"هزینه"	Decimal Number	Measure (Fact Attribute)	استفاده برای محاسبه هزینه کل، تحلیل سود و زیان
"تخفیف"	Whole Number	Measure (Fact Attribute)	استفاده در محاسبه میزان تخفیف، شناسایی اصلاحات مالی

استفاده در محاسبه مالیات کل و تحلیل مالی	Measure (Fact Attribute)	Whole Number	"مالیات"
ایجاد ارتباط با جدول Dim Subscribers و شمارش مشترکین	Key (Foreign Key)	Whole Number	"کد مشترک"
استفاده به‌عنوان فیلتر و Slicer (خانگی، اختصاصی و ...)	Dimension Attribute	Text	"نوع استفاده"
تحلیل و فیلتر جغرافیایی فروش	Dimension Attribute	Text	"شهر"
تحلیل منطقه‌ای و گزارش‌های مدیریتی	Dimension Attribute	Text	"استان"
دسته‌بندی سرویس‌ها و فیلتر تحلیلی	Dimension Attribute	Text	"گروه"
بررسی وضعیت سرویس و فیلتر رکوردها	Dimension Attribute	Text	"وضعیت"
تفکیک نوع مشتری در تحلیل‌ها	Dimension Attribute	Text	"حقیقی/حقوقی"
تحلیل کانال فروش (حضوری / آنلاین)	Dimension Attribute	Text	"نحوه خرید"
تحلیل عملکرد کاربران سیستم (در صورت نیاز)	Administrative Attribute	Text	"کاربر ایجاد فاکتور"
شناسایی یکتای فاکتور و شمارش تراکنش‌ها	Degenerate Dimension	Whole Number	"فاکتور ثبت"
تفکیک سرویس‌ها در گزارش‌ها	Dimension Attribute	Text	"نوع سرویس"
تحلیل نوع خدمت ارائه‌شده	Dimension Attribute	Text	"نوع خدمات"
تحلیل رفتار مشتریان جدید و تمدیدی	Dimension Attribute	Text	"خرید اولیه/تمدید"
تحلیل زمانی، روند فروش و فیلتر تاریخ	Date Attribute	Text	"تاریخ ثبت"
تحلیل عملکرد پکیج‌ها و مقایسه فروش	Dimension Attribute	Text	"پکیج"

1.5.1 هدف این بخش چیست در این مرحله تو: KPI جدید اختراع نمی‌کنی فقط KPIهای واقعاً مهم را انتخاب و شفاف می‌کنی بین داده و تصمیم کسب‌وکار پل می‌زنی چطور KPI ها را شناسایی کنیم؟ (روش حرفه‌ای ولی ساده) از داده شروع نکن؛ از «سوال» شروع کن از خودت بپرس: مدیر با این داده چه چیزی می‌خواهد بداند؟ چه تصمیمی قرار است بگیرد؟ مثال : فروش کم شده؟ KPI درآمد مشتری ریزش دارد؟ Churn Rate عملکرد منطقه‌ای؟ Revenue by Region هر KPI باید حداقل ۳ شرط را داشته باشد: قابل اندازه‌گیری قابل اقدام (Actionable) قابل توضیح در یک جمله و بعد از آن این جدول زیر را پر کن

KPI Name	Definition	Why Important	Decision
Total Revenue	مجموع مبلغ خالص تمامی فاکتورهای ثبت‌شده در بازه زمانی مشخص	شاخص اصلی سلامت مالی و عملکرد کلی فروش	ارزیابی رشد یا افت فروش و برنامه‌ریزی مالی
Total Cost	مجموع هزینه‌های ثبت‌شده مربوط به فروش	بررسی میزان هزینه‌کرد در کنار درآمد	کنترل هزینه‌ها و بهینه‌سازی ساختار مالی
Net Profit	اختلاف بین درآمد کل و هزینه کل	نشان‌دهنده سودآوری واقعی کسب‌وکار	تصمیم‌گیری درباره ادامه یا اصلاح استراتژی فروش
Number of Invoices	تعداد فاکتورهای صادر شده	سنجش حجم فعالیت فروش	ارزیابی عملکرد فروش مستقل از مبلغ
Active Customers	تعداد مشترکینی که حداقل یک فاکتور دارند	نشان‌دهنده پایگاه مشتریان فعال	تمرکز بر حفظ یا افزایش مشتریان فعال
Revenue per Customer	میانگین درآمد حاصل از هر مشترک	سنجش ارزش اقتصادی هر مشتری	تصمیم‌گیری درباره قیمت‌گذاری و پکیج‌ها
Revenue by Region	مجموع درآمد به تفکیک استان یا شهر	مقایسه عملکرد جغرافیایی	تمرکز بازاریابی بر مناطق ضعیف یا قوی
Renewal Ratio	نسبت فاکتورهای تمدید به کل فاکتورها	سنجش وفاداری مشتریان	بهبود سیاست‌های نگهداشت مشتری
Online Purchase Share	سهم خریدهای آنلاین از کل خریدها	بررسی عملکرد کانال فروش آنلاین	تصمیم برای توسعه یا اصلاح کانال فروش

1.6.1 هدف اصلی این تحلیل چیست؟ چه تصمیماتی قرار است بر اساس نتایج تحلیل گرفته شود؟ مشکل یا فرصت اصلی که قرار است داده‌ها برای آن پاسخ دهند چیست؟ چه معیارهایی موفقیت تحلیل را مشخص می‌کنند؟ داده‌های موجود تا چه اندازه برای رسیدن به هدف کافی هستند؟ آیا تحلیل باید به صورت توصیفی، تشخیصی، پیش‌بینی یا تجویزی باشد؟

هدف اصلی از انجام این تحلیل، تبدیل داده‌های خام فروش و اطلاعات مشترکین به بینش‌های قابل استفاده برای تصمیم‌گیری مدیریتی است، به‌گونه‌ای که وضعیت فعلی کسب‌وکار به‌صورت شفاف قابل مشاهده بوده و نقاط ضعف، قوت، فرصت‌ها و مسائل پنهان در عملکرد فروش و رفتار مشترکین شناسایی شوند. این تحلیل به مدیران کمک می‌کند تا درک دقیقی از میزان درآمد، سودآوری، رفتار خرید مشتریان، عملکرد پکیج‌ها، کانال‌های فروش و توزیع جغرافیایی درآمد به دست آورند و بر اساس آن تصمیمات آگاهانه‌تری اتخاذ کنند. تصمیماتی که قرار است بر مبنای نتایج این تحلیل گرفته شوند شامل اصلاح استراتژی فروش، تمرکز بر پکیج‌ها یا سرویس‌های سودآورتر، بهینه‌سازی کانال‌های فروش مانند خرید آنلاین و حضوری، افزایش نرخ تمدید مشترکین و تخصیص منابع بازاریابی به مناطق با عملکرد ضعیف‌تر است.

مسئله یا فرصت اصلی که این داده‌ها قرار است به آن پاسخ دهند، ارزیابی سلامت مالی و عملیاتی کسب‌وکار و شناسایی دلایل تغییرات درآمد و رفتار مشترکین در بازه زمانی مورد بررسی است. داده‌ها این امکان را فراهم می‌کنند که مشخص شود آیا کاهش یا افزایش فروش ناشی از کاهش تعداد مشترکین فعال است یا کاهش درآمد حاصل از هر مشترک، و همچنین آیا فرصت‌هایی برای افزایش درآمد از طریق بهبود نرخ تمدید، تغییر پکیج‌ها یا تمرکز بر مناطق جغرافیایی خاص وجود دارد.

معیارهای موفقیت این تحلیل شامل شفاف بودن شاخص‌های کلیدی عملکرد، قابلیت پاسخ‌گویی داشبورد به سوالات مدیریتی تعریف‌شده، امکان مقایسه عملکرد در ابعاد مختلف مانند زمان، منطقه و نوع مشتری، و توانایی استفاده عملی از نتایج تحلیل در تصمیم‌گیری‌های واقعی کسب‌وکار است. هرچه KPI ها ساده‌تر، قابل اندازه‌گیری‌تر و قابل اقدام‌تر باشند، تحلیل موفق‌تر تلقی می‌شود.

با توجه به ساختار داده‌های موجود، شامل جدول فروش با اطلاعات تراکنشی کامل و جدول مشترکین با داده‌های توصیفی گسترده، می‌توان گفت داده‌ها تا حد زیادی برای دستیابی به اهداف تحلیل کافی هستند، هرچند وجود برخی نواقص مانند مقادیر غیرمنطقی در ARPU، تاریخ‌های نامعتبر یا پراکندگی بالای داده‌های مالی، نیازمند پاک‌سازی و پیش‌پردازش است. با انجام این اصلاحات، داده‌ها توان پشتیبانی از تحلیل‌های مدیریتی معتبر را خواهند داشت.

از نظر نوع تحلیل، تمرکز اصلی این پروژه بر تحلیل توصیفی و تشخیصی است؛ به این معنا که ابتدا وضعیت فعلی کسب‌وکار به‌صورت عددی و بصری توصیف می‌شود و سپس دلایل احتمالی الگوها و تغییرات مشاهده‌شده بررسی می‌گردد. در این مرحله تحلیل پیش‌بینی یا تجویزی مدنظر نیست، اما نتایج به‌دست‌آمده می‌تواند در آینده مبنایی برای توسعه تحلیل‌های پیشرفته‌تر مانند پیش‌بینی رفتار مشترکین یا پیشنهاد اقدامات بهینه مدیریتی قرار گیرد.

1.7.1 هدف طراحی رویکرد بصری‌سازی داده هدف این مرحله، انتخاب روش مناسب نمایش داده‌ها است تا: اطلاعات به شکل ساده و قابل فهم ارائه شود الگوها، روندها و ناهنجاری‌ها به راحتی شناسایی شوند تصمیم‌گیرندگان بتوانند سریع و دقیق تصمیم بگیرند سوال‌های مهم برای تعیین هدف بصری‌سازی: داده‌ها شامل چه نوع متغیرهایی هستند؟ (عددی، دسته‌ای، زمانی، جغرافیایی...) مخاطب اصلی تحلیل چه کسانی هستند و چه سطحی از دانش دارند؟ آیا هدف شناسایی روند، توزیع، ارتباط یا مقایسه است؟ داده‌ها به چه میزان پیچیده هستند و چه میزان جزئیات باید نمایش داده شود؟ آیا لازم است اطلاعات به صورت تعامل‌پذیر (Interactive) ارائه شود؟

هدف از طراحی رویکرد بصری‌سازی داده‌ها در این پروژه، انتخاب مناسب‌ترین روش‌های نمایش اطلاعات به‌گونه‌ای است که داده‌های فروش و اطلاعات مشترکین به شکلی ساده، قابل فهم و در عین حال تحلیلی ارائه شوند و امکان شناسایی سریع الگوها، روندها و ناهنجاری‌ها برای تصمیم‌گیرندگان فراهم گردد. بررسی ساختار داده‌ها نشان می‌دهد که مجموعه داده‌های موجود شامل ترکیبی از متغیرهای عددی، دسته‌ای، زمانی و جغرافیایی است. متغیرهای عددی مانند مبلغ خالص، هزینه، تخفیف، مالیات و ARPU امکان تحلیل‌های کمی، محاسبه شاخص‌های کلیدی عملکرد و بررسی توزیع داده‌ها را فراهم می‌کنند. متغیرهای دسته‌ای مانند نوع سرویس، نوع خدمات، پکیج، وضعیت، نحوه خرید و حقیقی یا حقوقی بودن مشتری، برای مقایسه و تفکیک داده‌ها در تحلیل‌های چندبعدی مورد استفاده قرار می‌گیرند. همچنین وجود متغیرهای زمانی مانند تاریخ ثبت فاکتور، تاریخ ثبت نام و تاریخ آخرین فاکتور امکان تحلیل روندهای زمانی و بررسی تغییرات عملکرد در بازه‌های مختلف را فراهم می‌سازد. داده‌های جغرافیایی شامل استان و شهر نیز بستر لازم برای تحلیل‌های منطقه‌ای و مقایسه عملکرد جغرافیایی را ایجاد می‌کنند.

مخاطبان اصلی این تحلیل، مدیران و تصمیم‌گیرندگان کسب‌وکار در سطوح اجرایی و میانی هستند که انتظار دارند بدون ورود به جزئیات فنی پیچیده، تصویری شفاف از وضعیت فروش، درآمد و رفتار مشترکین دریافت کنند. بنابراین سطح دانش مخاطب به‌گونه‌ای است که بصری‌سازی‌ها باید ساده، گویا و مبتنی بر شاخص‌های کلیدی باشند و از نمایش بیش از حد جزئیات خام اجتناب شود، در حالی که امکان Drill Down در صورت نیاز حفظ گردد.

هدف اصلی بصری‌سازی در این پروژه، شناسایی و نمایش روندها، مقایسه عملکرد در ابعاد مختلف و تشخیص ناهنجاری‌ها است. نمودارهای روند برای نمایش تغییرات فروش و درآمد در طول زمان، نمودارهای مقایسه‌ای برای بررسی عملکرد پکیج‌ها، سرویس‌ها و مناطق جغرافیایی و نمودارهای توزیع برای شناسایی پراکندگی مقادیر مالی و داده‌های پرت، از اولویت‌های اصلی بصری‌سازی به شمار می‌روند. ارتباط بین متغیرها نیز به‌صورت محدود و در قالب مقایسه درآمد با نوع سرویس یا نوع مشتری بررسی می‌شود، بدون آنکه پیچیدگی بصری بیش از حد به داشبورد تحمیل گردد.

از نظر میزان پیچیدگی داده‌ها، اگرچه حجم رکوردها نسبتاً بالا است، اما ساختار داده‌ها منظم و قابل کنترل بوده و می‌توان با خلاصه‌سازی مناسب و استفاده از API ها، سطح جزئیات نمایش‌داده‌شده را در حد مورد نیاز مدیران نگه داشت. به همین دلیل، تمرکز بصری‌سازی بر نمایش خلاصه‌های مدیریتی در لایه اول و فراهم کردن امکان مشاهده جزئیات در لایه‌های بعدی قرار می‌گیرد.

با توجه به نیاز تصمیم‌گیرندگان به بررسی سناریوهای مختلف، مقایسه گروه‌ها و فیلتر کردن داده‌ها بر اساس زمان، منطقه یا نوع مشتری، ارائه اطلاعات به‌صورت تعاملی کاملاً ضروری است. استفاده از Slicer ها، فیلترهای پویا و قابلیت Drill Down در Power BI به کاربران اجازه می‌دهد بدون پیچیدگی فنی، داده‌ها را از زوایای مختلف بررسی کرده و تصمیمات سریع‌تر و دقیق‌تری اتخاذ کنند.

1.8.1 شناسایی موضوع اصلی کسب‌وکار ابتدا مشخص کنید سوال مربوط به چه بخش یا فرآیندی است: فروش بازاریابی مشتریان و رفتار آن‌ها عملیات و لجستیک مالی و سودآوری ۲. تعریف هدف تصمیم‌گیری مشخص کنید که چه تصمیمی قرار است بر اساس پاسخ گرفته شود. سوال باید به تصمیم‌گیری مرتبط باشد، نه صرفاً کشف داده‌ها. ۳. تعیین مخرج قابل اندازه‌گیری (Measurable Aspect) مشخص کنید که چه داده‌ها یا شاخص‌هایی برای پاسخ به سوال لازم است: تعداد فروش، درآمد، هزینه، سود تعداد مشتریان، میزان تعامل بازده سرمایه‌گذاری، نرخ بازگشت ۴. محدود کردن سوال به دامنه مشخص محدوده زمانی، جغرافیایی، یا گروه مشتریان را مشخص کنید. سوال باید عملی و واقع‌بینانه باشد. ۵. بررسی واضح و قابل فهم بودن سوال سوال باید ساده و قابل توضیح به مدیر یا ذی‌نفعان باشد. از اصطلاحات مبهم و کلی پرهیز کنید ؟
با بررسی ساختار داده‌های موجود شامل جدول فروش و جدول مشترکین، مشخص می‌شود که تمرکز اصلی این داده‌ها بر فرآیندهای فروش، رفتار مشتریان و مالی و سودآوری است. وجود اطلاعات تراکنشی مانند مبلغ خالص، هزینه، تخفیف، مالیات، نوع خرید و تمدید، در کنار اطلاعات توصیفی مشترکین مانند نوع مشتری، پکیج، وضعیت، ARPU و موقعیت جغرافیایی نشان می‌دهد که داده‌ها به‌طور مستقیم برای تحلیل عملکرد فروش و درآمدزایی، ارزیابی رفتار مشتریان و بررسی سودآوری کسب‌وکار طراحی شده‌اند و حوزه‌هایی مانند عملیات و لجستیک نقش کمرنگ‌تری در این مجموعه داده دارند. بنابراین موضوع اصلی کسب‌وکار در این تحلیل، ارزیابی عملکرد فروش و درآمد و نحوه تعامل و نگهداشت مشتریان است.

در گام بعد، هدف تصمیم‌گیری بر اساس داده‌ها مشخص می‌شود. تصمیماتی که قرار است با استفاده از پاسخ سوالات تحلیلی اتخاذ شوند شامل تشخیص وضعیت فروش در بازه زمانی مشخص، شناسایی کاهش یا افزایش درآمد، ارزیابی موفقیت تمدید اشتراک‌ها، تشخیص مناطق یا گروه‌های مشتریان با عملکرد ضعیف‌تر و تصمیم‌گیری درباره تمرکز منابع فروش و بازاریابی است. بنابراین سؤالات مطرح‌شده صرفاً برای کشف داده‌ها نیستند، بلکه مستقیماً به تصمیماتی مانند اصلاح پکیج‌ها، تغییر استراتژی فروش، تمرکز بر کانال‌های مؤثرتر یا بهبود نگهداشت مشتری منجر می‌شوند.

برای پاسخ‌گویی به این سؤالات، مخرج‌های قابل اندازه‌گیری به‌صورت شفاف از داده‌ها استخراج می‌شوند. شاخص‌هایی مانند مجموع درآمد، مجموع هزینه، سود خالص، تعداد فاکتورهای صادر شده، تعداد مشترکین فعال، درآمد به ازای هر مشترک، نسبت تمدید به خرید اولیه و درآمد به تفکیک استان یا شهر، مهم‌ترین معیارهای کمی هستند که داده‌های موجود امکان محاسبه دقیق آن‌ها را فراهم می‌کنند. این شاخص‌ها به‌طور مستقیم قابل اندازه‌گیری بوده و مبنای تعریف KPIها و تحلیل‌های مدیریتی قرار می‌گیرند.

در ادامه، سوالات تحلیلی به دامنه‌ای مشخص و عملی محدود می‌شوند تا قابلیت اجرا و استفاده در تصمیم‌گیری داشته باشند. داده‌های فروش دارای بازه زمانی مشخص و محدود هستند و امکان تحلیل در یک بازه زمانی کوتاه‌مدت وجود دارد. همچنین وجود

اطلاعات جغرافیایی مانند استان و شهر و دسته‌بندی مشتریان بر اساس نوع سرویس، پکیج و حقیقی یا حقوقی بودن، این امکان را فراهم می‌کند که سؤالات به گروه‌های مشخصی از مشتریان یا مناطق خاص محدود شوند و از کلی‌گویی پرهیز شود. این محدودسازی باعث می‌شود خروجی تحلیل قابل اقدام و متناسب با واقعیت‌های کسب‌وکار باشد.

در نهایت، سؤالات تعریف‌شده از نظر وضوح و قابل فهم بودن بررسی می‌شوند. با توجه به ماهیت داده‌ها، سوالاتی مانند «آیا فروش در بازه زمانی مورد بررسی کاهش یافته است؟»، «کدام استان‌ها بیشترین و کمترین درآمد را ایجاد کرده‌اند؟»، «سهم تمدیدها از کل فروش چقدر است؟» و «میانگین درآمد هر مشترک فعال چقدر است؟» سوالاتی ساده، شفاف و قابل توضیح برای مدیران و ذی‌نفعان هستند و از به‌کارگیری مفاهیم مبهم یا پیچیده اجتناب شده است. این ویژگی باعث می‌شود نتایج تحلیل به‌راحتی در جلسات مدیریتی ارائه شده و مبنای تصمیم‌گیری عملی قرار گیرند.

1.8.2 بررسی ریسک‌ها و محدودیت‌های داده؟ داده‌های موجود چقدر کامل و معتبر هستند؟ آیا داده‌ها به‌روز و دقیق هستند؟ آیا محدودیت‌های قانونی یا محرمانگی برای استفاده از داده‌ها وجود دارد؟ آیا حجم داده برای تحلیل کافی است؟ آیا داده‌ها به صورت یکپارچه و استاندارد ذخیره شده‌اند؟

با بررسی ساختار و محتوای داده‌های موجود، می‌توان گفت که از نظر پوشش فرآیند فروش و اطلاعات پایه مشتریان، داده‌ها تا حد قابل قبولی کامل هستند. وجود ستون‌هایی مرتبط با مبلغ خالص، هزینه، تخفیف، مالیات، نوع خرید، نوع سرویس، پکیج، وضعیت مشتری و اطلاعات جغرافیایی نشان می‌دهد که داده‌ها بخش عمده‌ای از جریان درآمدی و تعامل مشتری با کسب‌وکار را ثبت کرده‌اند. با این حال، نبود برخی فیلدهای کلیدی مانند شناسه یکتای تراکنش، تاریخ پایان سرویس، مدت اعتبار پکیج، یا وضعیت لغو اشتراک، باعث می‌شود داده‌ها از منظر تحلیلی «نسبتاً کامل» تلقی شوند نه کاملاً کامل. این محدودیت به‌ویژه در تحلیل‌هایی مانند محاسبه دقیق ریزش مشتری یا بررسی طول عمر مشتری اثرگذار است.

از نظر اعتبار داده‌ها، بخشی از ستون‌ها به‌صورت متنی و وابسته به ورود انسانی هستند، از جمله نوع استفاده، نحوه خرید، وضعیت، نوع خدمات و کاربر ایجاد فاکتور. این موضوع ریسک بروز خطاهای انسانی، عدم یکنواختی در نام‌گذاری و ثبت مقادیر مشابه با عناوین متفاوت را افزایش می‌دهد. در نتیجه، هرچند داده‌ها از نظر مفهومی معتبر هستند، اما از نظر اجرایی ممکن است دارای ناسازگاری‌هایی باشند که بر دقت دسته‌بندی و تجمیع داده‌ها اثر بگذارد.

در خصوص به‌روز بودن و دقت داده‌ها، ستون تاریخ ثبت تنها شاخص زمانی موجود است و به‌صورت متنی ذخیره شده است. این مسئله هم احتمال وجود فرمت‌های مختلف تاریخ و هم ریسک خطا در تحلیل‌های زمانی را افزایش می‌دهد. علاوه بر این، نبود ستون‌هایی مانند تاریخ آخرین بروزرسانی یا وضعیت فعال/غیرفعال جاری، ارزیابی دقیق به‌روز بودن واقعی رکوردها را دشوار می‌کند. بنابراین، داده‌ها برای تحلیل وضعیت گذشته مناسب هستند، اما برای تصمیم‌گیری‌های لحظه‌ای یا عملیاتی نیازمند احتیاط بیشتری می‌باشند.

از منظر محدودیت‌های قانونی و محرمانگی، داده‌ها شامل اطلاعات شناسایی مستقیم مانند نام، شماره تماس یا کد ملی مشتریان نیستند و عمدتاً از کد مشترک و ویژگی‌های توصیفی عمومی استفاده شده است. با این حال، وجود اطلاعاتی مانند شهر، استان، نوع مشتری (حقیقی/حقوقی) و الگوی خرید، داده‌ها را در دسته داده‌های نیمه‌حساس قرار می‌دهد. بنابراین استفاده از داده‌ها برای تحلیل داخلی و مدیریتی بلامانع است، اما انتشار عمومی داشبورد یا به‌اشتراک‌گذاری داده خام بدون حذف یا تجمیع اطلاعات مکانی و رفتاری می‌تواند ریسک نقض محرمانگی ایجاد کند.

در خصوص حجم داده، تعداد رکوردها برای انجام تحلیل‌های توصیفی، مقایسه‌ای و شناسایی الگوهای کلی فروش و رفتار مشتری کافی به نظر می‌رسد. حجم داده امکان تجمیع در سطوح مختلف مانند زمان، منطقه، نوع سرویس و نوع مشتری را فراهم می‌کند، هرچند برای تحلیل‌های پیش‌بینی دقیق یا مدل‌سازی پیشرفته ممکن است محدودیت‌هایی از نظر عمق تاریخی داده وجود داشته باشد.

در نهایت، از نظر یکپارچگی و استانداردسازی، داده‌ها به‌صورت متمرکز و در قالب جداول مشخص ذخیره شده‌اند، اما استفاده هم‌زمان از داده‌های عددی، متنی و تاریخ‌های غیرساخت‌یافته نشان می‌دهد که استانداردسازی کامل در سطح ذخیره‌سازی رعایت نشده است. این موضوع نیاز به پیش‌پردازش، یکسان‌سازی مقادیر متنی و تبدیل تاریخ‌ها به فرمت استاندارد را پیش از تحلیل الزامی می‌کند.

در مجموع، داده‌های موجود برای تحلیل‌های مدیریتی در سطح توصیفی و تشخیصی مناسب و قابل استفاده هستند، اما محدودیت‌هایی در کامل بودن، استانداردسازی، دقت زمانی و ملاحظات محرمانگی دارند که باید در تفسیر نتایج تحلیل و ارائه داشبورد نهایی به‌صورت شفاف مدنظر قرار گیرند.

1.8.3 تعیین نوع گزارش مناسب برای مخاطبان مختلف - مخاطب گزارش کیست؟ (مدیرعامل، تیم بازاریابی، تیم عملیاتی...)
مخاطب به چه جزئیاتی نیاز دارد؟ (خلاصه، نمودار، داده خام...) آیا گزارش باید تعاملی باشد یا ثابت؟ چه فرمت ارائه‌ای برای مخاطب راحت‌تر است؟ (داشبورد، PDF، نمودار، پاورپوینت) چه شاخص‌ها یا KPI هایی برای مخاطب اهمیت دارند؟
در این مرحله، با توجه به ساختار داده‌های موجود شامل جدول Fact_Sales و Dim_Subscribers، نوع گزارش و نحوه ارائه اطلاعات بر اساس نیاز مخاطبان مختلف کسب‌وکار تعیین می‌شود. مخاطب اصلی گزارش‌ها شامل مدیران ارشد، تیم بازاریابی و فروش، تیم عملیاتی و پشتیبانی و همچنین تحلیل‌گران داده است. برای مدیرعامل و مدیران ارشد، گزارش باید در سطح خلاصه و کلان ارائه شود و تمرکز آن بر شاخص‌های کلیدی عملکرد مانند مبلغ خالص فروش، روند درآمد در طول زمان بر اساس تاریخ ثبت، سودآوری کلی، توزیع درآمد به تفکیک استان و نسبت خرید اولیه به تمدید باشد. این گزارش ترجیحاً به‌صورت داشبورد تعاملی در Power BI طراحی می‌شود تا امکان فیلتر زمانی وجود داشته باشد و در عین حال یک خروجی خلاصه به‌صورت PDF یا اسلاید مدیریتی برای جلسات تصمیم‌گیری در دسترس باشد. برای تیم بازاریابی و فروش، گزارش نیازمند جزئیات بیش‌تری است و باید امکان بررسی عملکرد فروش به تفکیک نوع سرویس، پکیج، نوع مشتری، نحوه خرید و منطقه جغرافیایی فراهم شود. شاخص‌هایی مانند درآمد هر پکیج، میانگین مبلغ فاکتور، میزان و اثر تخفیف‌ها، نسبت تمدید به خرید اولیه و شاخص ARPU از جدول مشترکین در این گزارش اهمیت بالایی دارند و قالب مناسب آن داشبورد کاملاً تعاملی Power BI با قابلیت Drill-down و فیلترهای متنوع است. در مورد تیم عملیاتی و پشتیبانی، تمرکز گزارش بر فرآیندها و وضعیت‌ها است و شاخص‌هایی مانند تعداد فاکتورهای ثبت‌شده، توزیع وضعیت سرویس‌ها، عملکرد کاربران ثبت‌کننده فاکتور و نوع خدمات ارائه‌شده اهمیت دارد؛ در این حالت گزارش باید ساده، کاربردی و بیشتر مبتنی بر جداول و نمودارهای توزیعی باشد. در نهایت، برای تحلیل‌گران داده و تیم هوش تجاری، گزارش باید دسترسی کامل به داده‌های خام، مدل داده، رابطه Fact و Dimension و شاخص‌های کیفیت داده را فراهم کند تا امکان تحلیل عمیق، شناسایی ناهنجاری‌ها و توسعه Measure های تحلیلی با DAX وجود داشته باشد. این رویکرد چندسطحی در گزارش‌دهی باعث می‌شود هر گروه از مخاطبان، دقیقاً به اندازه نیاز خود اطلاعات دریافت کرده و تصمیم‌گیری سریع‌تر و دقیق‌تری در سطح کسب‌وکار انجام شود.

1.9.1 و در انتها با توجه به تحقیقات و داده های بدست آمده و جلسات با کارفرما یک گزارش کامل و اصولی برای ردمی گیت امده کرده که ساختار گزارش تا این قسمت را شرح دهد و حتما برای داده بیشتر به این قسمت هم ارجاع دهید و در نهایت کامیت بخش اول

2. مهندسی داده و مدل‌سازی

2.0.1 ضبط ویدیو از طول فرایند اجرایی در طول اجرای این مرحله، فرآیند کامل مهندسی داده و مدل‌سازی به صورت ویدیو ضبط شد. هدف از ضبط، مستندسازی و ارائه شفاف فرآیند برای انتشار آموزشی در فضای مجازی بود. برای ضبط ویدیو از نرم‌افزار Techsmith استفاده شده است. لازم به ذکر است که محتوای ضبط شده تنها شامل مراحل اجرایی و تغییرات داده‌ها است و هیچ اطلاعات حساس یا خصوصی منتشر نشده است.

2.1.1 چک‌لیست آماده‌سازی داده‌ها برای ETL با Power Query اول از همه - ۱. بررسی داده‌های اولیه اطمینان از دریافت تمام فایل‌ها و جداول مورد نیاز ؟ بررسی فرمت فایل‌ها (CSV, Excel, JSON، ...) ؟ بررسی تعداد ردیف و ستون‌ها با مستندات اولیه ؟ شناسایی ستون‌های کلیدی و مهم برای تحلیل ۲. پاک‌سازی داده‌ها (Data Cleaning) - در صورت مشاهده داده اشتباه در هدر اقدام به اصلاح آن ؟ حذف ردیف‌های خالی یا ناقص (در صورت مشاهده داده یا ردیف که اکثر داده های آن خالی هستند یا داده های مهم آن خالی هستند) ؟ اصلاح یا حذف مقادیر نامعتبر یا غیر منطقی (مثلاً تاریخ‌های آینده، اعداد منفی در فروش البته حتما این مشکل را با کارفرما در میان گذاشته با تایید آن اقدام به حذف آنها انجام شود) ؟ استانداردسازی نام ستون‌ها و داده‌ها (CamelCase یا snake_case) ؟ اصلاح فرمت تاریخ و زمان ؟ حذف فاصله‌ها یا کاراکترهای اضافی در متن (و معمولا این مشکل پیش می آید و ان صلاح حرف ک و ی عربی به فارسی هست) ۳. تبدیل داده‌ها (Data Transformation) ؟ تغییر نوع داده‌ها (Data Type) به مناسب (عدد، متن، تاریخ) ؟ ایجاد ستون‌های محاسباتی جدید در صورت نیاز (مثلاً سود = فروش – هزینه) ؟ ترکیب یا تفکیک ستون‌ها در صورت لزوم (Merge/Split Columns) ؟ پالایش و استانداردسازی مقادیر دسته‌ای یا گروه(مثلاً نام استان‌ها یا دسته‌بندی محصولات) ؟ حذف ستون‌های غیر ضروری برای تحلیل ۴. آماده‌سازی داده‌ها برای گزارش - اطمینان از یکپارچگی داده‌ها بین جداول ؟ اطمینان از عدم وجود تکراری‌ها (در این قسمت هم از کارفرما حتما پرسیده به این دلیل که در بعضی از داده داده تکرای موجود هست و این کامل طبیعی است) ؟ بررسی داده‌ها با نمونه واقعی گزارش‌ها یا KPI ها مستندسازی تغییرات اعمال شده روی داده‌ها (چه چیزی حذف، اصلاح یا اضافه شده) ۵. خروجی نهایی ETL - ذخیره جدول نهایی در قالب مناسب (CSV، Excel، یا اتصال مستقیم به Power BI) ؟ نامگذاری فایل‌ها با نسخه و تاریخ (مثلاً SalesData_Cleaned_v1_2026-02-01.csv) ؟ آماده کردن فایل برای مدل‌سازی داده و گزارش نهایی

ETL Preparation – Fact_Sales :

در فرآیند ETL جدول Fact_Sales، داده‌ها از فایل Excel منبع دریافت و شیت مربوط به فروش انتخاب شد. در گام نخست، هدرهای جدول اصلاح و به‌صورت استاندارد Promote شدند تا ساختار ستون‌ها برای تحلیل قابل استفاده باشد. سپس نوع داده ستون‌ها با توجه به ماهیت هر فیلد اصلاح شد؛ ستون‌های عددی مانند مبلغ خالص، هزینه، تخفیف، مالیات، فاکتور ثبت و کد مشترک به نوع عددی مناسب تبدیل شدند و ستون‌های متنی مانند نوع سرویس، پکیج، استان، شهر و وضعیت به‌صورت Text نگه داشته شدند. در ادامه، ستون‌های تاریخ شروع و پایان فاکتور که در تحلیل فعلی کاربرد نداشتند و از نظر کسب‌وکار نیز ضروری تشخیص داده نشدند، حذف شدند تا جدول Fact سبک‌تر و متمرکز بر تراکنش فروش باقی بماند. ستون «تاریخ ثبت» که شامل تاریخ و زمان به‌صورت ترکیبی بود، به دو ستون مجزا شامل «تاریخ ثبت» و «زمان تاریخ ثبت» تفکیک شد تا امکان تحلیل زمانی دقیق‌تر و اتصال به جدول تاریخ فراهم شود. در پایان این مرحله، جدول Fact_Sales به‌عنوان جدول تراکنشی اصلی با داده‌های پاک‌سازی‌شده، نوع داده استاندارد و بدون ستون‌های غیرضروری، آماده ورود به مرحله مدل‌سازی داده شد.

ETL Preparation – Dim_Subscribers

در فرآیند ETL جدول Dim_Subscribers، داده‌ها از همان فایل Excel و از شیت مربوط به مشترکین استخراج شدند. پس از Promote Header، نوع داده ستون‌ها بر اساس نقش تحلیلی آن‌ها اصلاح شد؛ ستون «کد مشترک» به‌عنوان کلید اصلی به نوع عددی تبدیل شد و سایر ستون‌ها شامل اطلاعات توصیفی، جمعیت‌شناختی و رفتاری مشترکین به‌صورت Text نگه داشته شدند. ستون ARPU به‌عنوان یک شاخص عددی مهم، به نوع عددی تبدیل شد تا در محاسبات تحلیلی قابل استفاده باشد. ستون‌های تاریخ مرتبط با چرخه عمر مشترک (مانند تاریخ ثبت نام، آخرین فاکتور و آخرین تغییرات) در این مرحله به‌صورت Text حفظ شدند تا در فاز بعدی و با کمک جدول تاریخ (Dim_Date) استانداردسازی و تبدیل شوند. هدف این مرحله، حفظ کامل اطلاعات توصیفی مشترکین و آماده‌سازی جدول

به‌عنوان Dimension اصلی برای اتصال به Fact_Sales از طریق کد مشترک بود.

ETL Preparation – Reload Table

جدول Reload به‌عنوان یک جدول کنترلی ایجاد شد که تنها شامل یک ستون با نام «Data Last Refreshed» است. این جدول به‌صورت داینامیک زمان آخرین Refresh داده‌ها را با استفاده از DateTime.LocalNow ثبت می‌کند. وجود این جدول امکان نمایش زمان آخرین به‌روزرسانی داده‌ها در داشبورد Power BI را فراهم کرده و از منظر کنترل کیفیت داده و اعتماد کاربران نهایی به گزارش‌ها اهمیت بالایی دارد.

ETL Preparation – Dim_Measure

جدول Dim_Measure به‌عنوان یک جدول ساختاری برای نگهداری و سازمان‌دهی Measure ها در مدل داده ایجاد شده است. این جدول به‌صورت فیزیکی داده تحلیلی ندارد و نقش آن صرفاً کمک به ساختاردهی، خوانایی مدل و مدیریت بهتر Measure های DAX در Power BI است. استفاده از Dim_Measure باعث جداسازی منطق محاسباتی از داده‌های خام شده و نگهداری مدل تحلیلی را ساده‌تر می‌کند.

ETL Preparation – Dim_Date

جدول Dim_Date به‌عنوان Dimension تاریخ مستقل طراحی شده و شامل مجموعه‌ای کامل از تاریخ‌های شمسی و میلادی است. این جدول نقش کلیدی در تبدیل تاریخ‌های متنی موجود در Fact_Sales و Dim_Subscribers، استخراج سال، ماه، روز، نام ماه و تحلیل‌های زمانی ایفا می‌کند. وجود Dim_Date امکان تحلیل روندها، مقایسه دوره‌ای و گزارش‌گیری استاندارد زمانی را فراهم کرده و وابستگی مستقیم به فرمت‌های تاریخ ناهمگون در داده‌های خام را از بین می‌برد.

2.2.1 در فرایند ETL، به‌ویژه هنگام کار با Power Query، معمولاً لازم است در طول پروژه چندین بار به مراحل قبلی بازگردیم و تغییرات جدیدی روی داده‌ها اعمال کنیم. این بازگشت‌ها می‌تواند به دلایل مختلفی مانند تغییر نیاز تحلیل، درخواست کارفرما، یا کشف خطا در داده‌ها انجام شود. به همین دلیل، مستندسازی فرایند ETL نقش بسیار مهمی در شفافیت، پاسخ‌گویی و جلوگیری از سوء تفاهم دارد. در این مرحله، تمرکز اصلی بر ثبت دقیق تمام مراحل انجام‌شده از ابتدا تا انتهای پروژه و همچنین تغییراتی است که در طول زمان روی داده‌ها اعمال شده‌اند. مستندسازی به ما کمک می‌کند: بدانیم چه تغییری، در چه زمانی و با چه هدفی انجام شده است در صورت نیاز بتوانیم مراحل قبلی را بازسازی یا اصلاح کنیم در مواجهه با کارفرما یا تیم پروژه، دلیل تصمیمات فنی را به‌صورت شفاف ارائه دهیم نکته تجربی (درس آموخته‌شده از پروژه واقعی) در یکی از پروژه‌ها، بخشی از داده که در تحلیل اولیه مورد استفاده قرار نگرفته بود، از Power Query حذف شد. در ادامه پروژه، مشخص شد که برای یک به‌روزرسانی جدید، نیاز به همان بخش حذف‌شده وجود دارد. از آن‌جایی که آن تغییر به‌صورت مکتوب و مستند ثبت نشده بود، کارفرما مسئولیت این مشکل را متوجه مجری پروژه دانست؛ در حالی که پیش از حذف داده، هماهنگی شفاهی انجام شده بود اما به‌دلیل فراموشی، قابل استناد نبود.

مرحله	نام جدول	نام اقدام	نوع عملیات	توضیح تغییر اعمال‌شده	دلیل انجام
1	Fact_Sales	دریافت داده از فایل اکسل	Extract	بارگذاری فایل اکسل «Sales Data.xls» و انتخاب شیت «فروش» به‌عنوان منبع داده خام	دسترسی به داده‌های اولیه فروش
2	Fact_Sales	ارتقاء سرفصل ستون‌ها	Transform	تبدیل اولین سطر داده به نام ستون‌ها با استفاده از Promote Headers	استانداردسازی ساختار جدول و آماده‌سازی برای تبدیل نوع داده
3	Fact_Sales	تعیین نوع داده ستون‌ها	Transform	تعیین نوع داده مناسب برای ستون‌های عددی، متنی و تاریخ (Int64، Number، Text)	جلوگیری از خطا در محاسبات، فیلترها و بصری‌سازی
4	Fact_Sales	حذف ستون‌های غیرضروری تاریخ	Clean	حذف ستون‌های «تاریخ شروع فاکتور» و «تاریخ پایان فاکتور»	عدم نیاز به این ستون‌ها در تحلیل فعلی و کاهش پیچیدگی مدل
5	Fact_Sales	تفکیک تاریخ و زمان ثبت	Transform	جداسازی ستون «تاریخ ثبت» به دو ستون تاریخ و زمان بر اساس جداکننده فاصله	امکان تحلیل دقیق‌تر زمانی (تاریخ و ساعت)
6	Fact_Sales	اصلاح نوع داده زمان	Transform	تبدیل بخش زمان به نوع Time	سازگاری با توابع زمانی Power BI
7	Fact_Sales	تغییر نام ستون‌ها	Transform	تغییر نام ستون‌ها به «تاریخ ثبت» و «زمان تاریخ ثبت»	افزایش خوانایی و شفافیت معنایی ستون‌ها
1	Dim_Subscribers	دریافت داده از فایل اکسل	Extract	بارگذاری فایل اکسل و انتخاب شیت «مشترکین» به‌عنوان منبع داده	دسترسی به اطلاعات پایه مشترکین
2	Dim_Subscribers	ارتقاء سرفصل ستون‌ها	Transform	تبدیل اولین سطر به عنوان نام ستون‌ها	آماده‌سازی جدول برای پردازش‌های بعدی
3	Dim_Subscribers	تعیین نوع داده ستون‌ها	Transform	تعیین نوع داده مناسب برای ستون‌های عددی، متنی و تاریخ	تضمین صحت تحلیل‌های آماری و ارتباط با جدول Fact
1	Dim_Date	دریافت داده از فایل اکسل	Extract	اضافه کردن جدول تاریخ شامل تاریخ‌های میلادی و شمسی	پشتیبانی از تحلیل‌های زمانی چند تقویمی
2	Dim_Date	تعیین نوع داده ستون‌ها	Transform	تعیین نوع داده مناسب برای ستون‌های عددی، متنی و تاریخ	تضمین صحت تحلیل‌های آماری و ارتباط با جدول Fact
1	Dim_Measure	ایجاد جدول معیارها	Create	ایجاد جدول مستقل برای نگهداری Measure ها	تفکیک منطق محاسباتی از داده خام
2	Dim_Measure	سازمان‌دهی Measure ها	Transform	دسته‌بندی و نام‌گذاری استاندارد Measure ها	افزایش خوانایی مدل و سهولت نگهداری

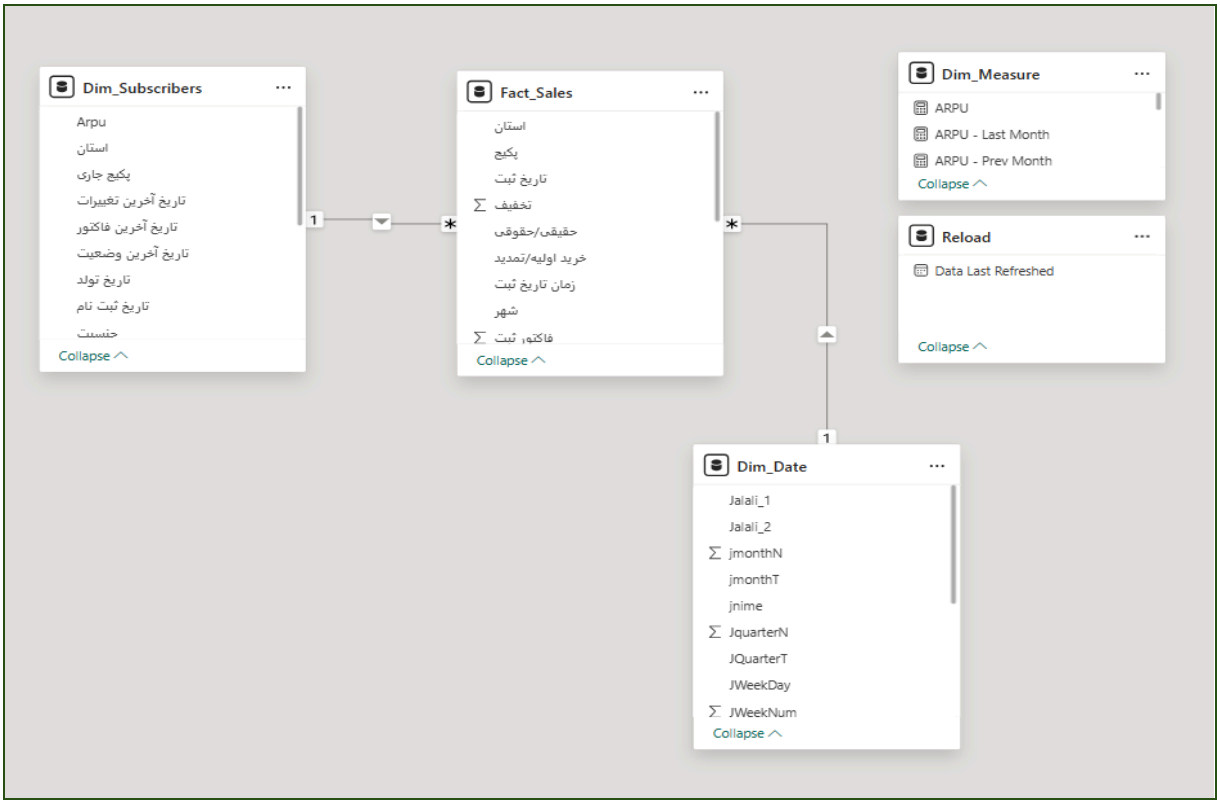
2.3.1 در این مرحله، تمرکز بر انتخاب مدل مناسب داده برای بهبود خوانایی، سادگی تحلیل و عملکرد گزارش‌ها بوده است از چه مدل داده‌ای استفاده شده است؟ چرا این مدل برای این پروژه انتخاب شده است؟

در این مرحله، برای طراحی مدل داده از مدل ستاره‌ای (Star Schema) استفاده شده است. در این مدل، یک جدول مرکزی از نوع Fact با نام Fact_Sales قرار دارد که شامل رویدادهای اصلی کسب‌وکار (تراکنش‌های فروش) و مقادیر عددی قابل تحلیل مانند مبلغ خالص، هزینه، تخفیف و مالیات است. در اطراف این جدول، جداول Dimension شامل Dim_Subscribers (اطلاعات توصیفی مشترکین)، Dim_Date (مدیریت تاریخ‌های شمسی و میلادی و استخراج روز، ماه و سال) و Dim_Measure (استانداردسازی و مدیریت شاخص‌ها و معیارها) قرار گرفته‌اند.

انتخاب مدل ستاره‌ای برای این پروژه به دلایل زیر انجام شده است:

اول، خوانایی و سادگی تحلیل؛ این مدل ساختاری شفاف دارد و برای کاربران تجاری و تحلیل‌گران به‌راحتی قابل درک است. دوم، بهبود عملکرد Power BI؛ مدل ستاره‌ای با کاهش پیچیدگی روابط و استفاده از Join های ساده، باعث افزایش سرعت پردازش و بهینه‌سازی Query ها می‌شود. سوم، تطابق کامل با ماهیت داده‌ها؛ داده‌های فروش ماهیت تراکنشی دارند و به‌طور طبیعی در قالب Fact Table مدل می‌شوند، در حالی که اطلاعات مشترکین، زمان و ویژگی‌های توصیفی بهترین گزینه برای Dimension Table هستند. چهارم، مقیاس‌پذیری و توسعه‌پذیری؛ در صورت اضافه شدن KPI ها، ابعاد جدید (مانند Dim_Product یا Dim_Channel) یا داده‌های آینده، این مدل بدون تغییرات اساسی قابل گسترش است.

در مجموع، استفاده از مدل ستاره‌ای باعث شده است که داشبورد نهایی سریع، قابل نگهداری، قابل توسعه و کاملاً منطبق با نیازهای تصمیم‌گیری مدیریتی باشد.



2.4.1 در این مرحله، پس از طراحی مدل داده با رویکرد Star Schema، ارتباط بین جداول به‌صورت کنترل‌شده، شفاف و بهینه تعریف شده است تا از صحت تحلیل‌ها و عملکرد مناسب مدل اطمینان حاصل شود چه ارتباط‌های ایجاد شده در مدل داده ؟ چرا این نوع

ارتباط‌ها انتخاب شده‌اند؟

پس از طراحی مدل داده با رویکرد Star Schema، ارتباط بین جداول به‌صورت کنترل‌شده، یک‌جهته و مبتنی بر کلیدهای منطقی کسب‌وکار تعریف شده است تا هم از صحت تحلیل‌ها اطمینان حاصل شود و هم عملکرد مدل در Power BI بهینه بماند.

در این مدل، جدول Fact_Sales به‌عنوان جدول مرکزی (Fact) در نظر گرفته شده و تمامی ارتباط‌ها از جداول Dimension به این جدول برقرار شده‌اند.

ارتباط اول بین Dim_Subscribers و Fact_Sales از طریق ستون کد مشترک ایجاد شده است. این ارتباط از نوع یک به چند (One-to-Many) است؛ به این معنا که هر مشترک می‌تواند چندین رکورد فروش داشته باشد، اما هر رکورد فروش فقط به یک مشترک تعلق دارد. جهت فیلترگذاری از Dimension به Fact تعریف شده تا امکان Slice و Filter داده‌های فروش بر اساس ویژگی‌های مشترکین (مانند استان، شهر، نوع مشتری و وضعیت) فراهم شود و از ایجاد ابهام در محاسبات جلوگیری گردد.

ارتباط دوم بین Dim_Date و Fact_Sales از طریق ستون تاریخ ثبت برقرار شده است. این ارتباط نیز از نوع یک به چند بوده و امکان تحلیل‌های زمانی مانند بررسی روند فروش، مقایسه ماهانه، فصلی و سالانه (بر اساس تقویم شمسی و میلادی) را فراهم می‌کند. استفاده از Dim_Date باعث شده تمام تحلیل‌های زمانی استاندارد، قابل توسعه و مستقل از داده خام باشند و محاسبات Time Intelligence با دقت بالاتری انجام شوند.

جدول Dim_Measure به‌صورت مستقیم رابطه فیزیکی با Fact_Sales ندارد و به‌عنوان یک Disconnected Table استفاده شده است. هدف از این جدول، مدیریت و انتخاب پویای شاخص‌ها (مانند ARPU، ARPU ماه قبل و ماه گذشته) در سطح لایه محاسباتی (DAX) است. این رویکرد باعث افزایش انعطاف‌پذیری گزارش و جلوگیری از پیچیدگی غیرضروری در مدل روابط شده است.

جدول Reload نیز یک جدول کمکی مستقل است که صرفاً برای نمایش آخرین زمان به‌روزرسانی داده‌ها در داشبورد استفاده می‌شود و به‌دلیل ماهیت نمایشی، هیچ ارتباطی با سایر جداول ندارد.

2.5.1 در این مرحله، هدف طراحی یک Measure Table استاندارد و تعریف Measure های اولیه مورد نیاز تحلیل است

Measure Table به‌عنوان یک جدول منطقی برای سازمان‌دهی، خوانایی و مدیریت بهتر محاسبات در مدل داده استفاده می‌شود الگوی تصمیم‌گیری برای ساخت Measure قبل از اضافه کردن هر Measure، این سوال‌ها بررسی می‌شوند: آیا این Measure در بیش از یک گزارش استفاده خواهد شد؟ آیا منطق محاسبه آن تکراری است؟ آیا برای تصمیم‌گیری کارفرما یا تحلیل ضروری است؟ اگر پاسخ مثبت باشد، Measure در Measure Table تعریف می‌شود. وقتی تعداد Measure ها یا داده‌های عددی زیاد است استفاده از Python باعث می‌شود فرآیند مهندسی داده سریع‌تر، مقیاس‌پذیرتر و قابل کنترل‌تر شود.

در لایه‌ی اول، Total Sales نمایانگر حجم کل فروش شامل مبلغ خالص، هزینه‌ها و مالیات است و به‌عنوان شاخص اصلی عملکرد فروش در نظر گرفته می‌شود. تغییرات هفتگی فروش از طریق WoW Change % بررسی شده که نشان می‌دهد فروش نسبت به هفته گذشته چه میزان رشد یا افت داشته و متن تفسیری آن در WoW Text به‌صورت قابل فهم ارائه شده است.

در بعد سودآوری، Total Profit میزان درآمد خالص کسب‌وکار را نشان می‌دهد و در کنار آن Total Cost و Total Tax امکان تحلیل دقیق‌تر ساختار مالی و فشار هزینه‌ای را فراهم می‌کنند. همچنین Total Discount برای بررسی سیاست‌های تخفیف و تأثیر آن بر فروش در دسترس است.

از منظر مشتری، Total Subscribers تعداد مشترکین فعال را مشخص می‌کند و شاخص ARPU (میانگین درآمد به ازای هر مشترک) کیفیت درآمدزایی از هر مشتری را نشان می‌دهد. این شاخص کمک می‌کند تشخیص دهیم رشد فروش ناشی از افزایش تعداد مشترکین است یا افزایش ارزش هر مشترک.

برای تحلیل روند ماهانه، تاریخ آخرین فاکتور ثبت‌شده به میلادی محاسبه شده و بازه‌های ماه جاری و ماه قبل به‌صورت پویا تعیین شده‌اند. بر این اساس، فروش، سود، تعداد مشترکین و ARPU برای ماه اخیر و ماه قبل محاسبه شده و نرخ تغییر ماه‌به‌ماه (MoM %) برای هرکدام ارائه شده است. این مقایسه‌ها امکان تشخیص روندهای پایدار رشد یا هشدار افت عملکرد را فراهم می‌کنند.

در نهایت، Number of Invoices به‌عنوان شاخص حجم فعالیت فروش، تعداد کل تراکنش‌ها را نشان می‌دهد و به تفکیک اثر «تعداد خرید» از «مبلغ خرید» کمک می‌کند.

به‌طور کلی، این مجموعه Measure ها یک زنجیره‌ی کامل از فروش → سود → مشتری → روند زمانی را پوشش می‌دهد و مبنای تصمیم‌گیری‌های مدیریتی در حوزه‌ی قیمت‌گذاری، کنترل هزینه، رشد فروش و نگهداشت مشترکین قرار می‌گیرد.

2.6.1 و در انتها با توجه به تحقیقات و داده های بدست آمده و جلسات با کارفرما یک گزارش کامل و اصولی برای ردمی گیت امده کرده که ساختار گزارش تا این قسمت را شرح دهد و حتما برای داده بیشتر به این قسمت هم ارجاع دهید و در نهایت کامیت بخش دوم

3. طراحی تجربه کاربری و داشبورد

3.0.1 ضبط ویدیو از طول فرایند اجرایی در طول اجرای این مرحله، فرآیند کامل مهندسی داده و مدل‌سازی به صورت ویدیو ضبط شد. هدف از ضبط، مستندسازی و ارائه شفاف فرآیند برای انتشار آموزشی در فضای مجازی بود. برای ضبط ویدیو از نرم‌افزار Techsmith استفاده شده است. لازم به ذکر است که محتوای ضبط شده تنها شامل مراحل اجرایی و تغییرات داده‌ها است و هیچ اطلاعات حساس یا خصوصی منتشر نشده است.

3.1.1 فهمیدن این نیست که «چه داده‌ای داریم» بلکه این است که «چه تصمیمی قرار است گرفته شود» Outcome (نتیجه)کارفرما بعد از دیدن داشبورد چه کاری باید انجام دهد؟ Decision (تصمیم) برای رسیدن به آن نتیجه، چه تصمیمی باید بگیرد؟ Question 3 (سؤال)برای گرفتن آن تصمیم، چه چیزی باید بداند؟ (سؤال‌ها همیشه آخر می‌آیند، نه اول) بخش دوم: روش استخراج سؤال‌ها (Step-by-Step) گام ۱: ممنوعیت سؤال‌های دیتایی در ابتدا حق نداری این‌ها را بپرسی: چه ستون‌هایی داریم؟ چه گزارشی می‌خواهید؟ چه KPI هایی لازم است؟ گام ۲: سؤال‌های «تحریک تصمیم» کشف الگوهای تکرارشونده (خیلی مهم) چرا نیاز به تحلیل علت - کجا بعد مکانی - کی بعد زمانی - کدوم بهتره مقایسه - همیشه یا هیچ وقت استثنا گام ۴: تبدیل حرف به سؤال تحلیلی بخش سوم: چارچوب گفتگو با کارفرما (قابل استفاده در پروژه) فاز ۱ — تصویر کلی «موفقیت از نظر شما یعنی چی؟» «بدترین سناریو چیه؟» فاز ۲ — کنترل «کجا بیشترین ابهام رو دارید؟» «الان تصمیم‌ها بیشتر بر چه اساسی گرفته می‌شن؟» فاز ۳ — اولویت «اگر فقط ۳ چیز رو بشه دید، اون‌ها چی هستن؟» نمونه از سوالات : چه تعداد از محصولات، 80٪ سود را ایجاد می‌کنند؟ سود تجمعی این ماه در مقایسه با آخرین ماه کامل چه تغییری کرده؟ کدام محصولات به‌صورت مداوم زیان‌ده بوده‌اند، نه مقطعی؟ یا رشد سود ناشی از افزایش حجم است یا کاهش هزینه؟ بزرگ‌ترین منبع ریسک سوددهی در حال حاضر کجاست؟ کدام دسته محصولات بیشترین تغییر رفتار (رشد یا افت) را نسبت به دوره قبل داشته‌اند؟

3.2.1 هدف این مرحله صرفاً مرتب‌سازی سوال‌ها نیست؛ بلکه تبدیل سوال‌ها به ساختار داشبورد و صفحات گزارش است. چرا دسته‌بندی سوالات مهم است؟ جلوگیری از پراکندگی بصری داشبورد مشخص شدن صفحات گزارش قبل از طراحی تفکیک تصمیم‌ها (نه داده‌ها) کمک به اولویت‌بندی KPI ها اصل پایه دسته‌بندی سوال‌ها بر اساس «موضوع تصمیم‌گیری» گروه‌بندی می‌شوند، نه بر اساس ستون یا جدول داده دسته‌های مرجع (قابل استفاده در اکثر پروژه‌ها) الف) Financial / Performance - ب) Product / Service - ج) Customer / Market - د) Operations / Process - هـ) Resources / People - اولویت‌بندی سوال‌ها داخل هر دسته Core: سوال‌های تصمیم‌ساز → صفحه اصلی - Secondary: سوال‌های تحلیلی → Drill-down - Optional: سوال‌های تکمیلی → Tooltip / Appendix پس از استخراج سوالات کلیدی، آن‌ها بر اساس نوع تصمیم‌گیری و اولویت تحلیلی دسته‌بندی شدند تا ساختار صفحات داشبورد به‌صورت هدفمند و قابل توسعه طراحی شود.

3.3.1 داشبورد بر اساس یک سناریوی داستان‌پردازی مرحله‌به‌مرحله طراحی شد که ابتدا تصویر کلی عملکرد را نمایش می‌دهد، سپس با تمرکز بر بخش‌های بحرانی، الگوهای کلیدی را آشکار کرده و در نهایت مسیر تصمیم‌گیری را برای مخاطب مشخص می‌کند (با توجه به توضیحات در قسمت پیاده سازی داشبورد این اصول را در آن پیاده سازی کن)

3.4.1 ساخت یک تمپلیت بصری که هویت برند را منتقل کند، تمرکز تصمیم را هدایت کند و در همه پروژه‌ها قابل تکرار باشد اول: انتخاب نوع تمپلیت (قبل از رنگ!) قبل از رنگ، باید سبک تمپلیت را مشخص کنی. ۳ سبک اصلی داریم 1. Executive / Decision-Oriented فضای سفید زیاد KPIهای درشت رنگ کم، معنی‌دار 2. Analytical / Deep-Dive نمودارهای بیشتر رنگ‌های خنثی + تاکید محدود جزئیات قابل جزئیات قابل Drill - کاربرد: کشف الگو، علت‌یابی 3. Operational / Monitoring هشدار رنگی Threshold مشخص وضعیت‌محور کاربرد: کنترل عملیات استخراج هویت برند (حتی اگر برند رسمی نداری) - Primary Color - Secondary Color - Accent Color - Font - سیستم رنگ‌بندی حرفه‌ای (نه انتخاب رنگ دلخواه) نگ = معنا، نه تزئین لایه‌بندی رنگ‌ها لایه 1: رنگ پایه Background - Grid - Card لایه 2: رنگ برند KPI اصلی - عنوان‌ها - عناصر کلیدی لایه 3: رنگ تاکید Highlight - انتخاب‌ها - Focus لایه 4: رنگ معنادار مثبت سبز - منفی قرمز - هشدار نارنجی - خنثی خاکستری، رنگ‌بندی داده‌ها و نمودارها (خیلی مهم) قانون 60–30–10 : 60٪ خنثی - 30٪ برند - 10٪ تاکید / هشدار، سطح‌بندی بصری (Visual Hierarchy) : مهمترین رنگ اندازه - متوسط اندازه - کم اهمیت شفافیت کمتر پیاده‌سازی تمپلیت در Power BI (استاندارد) : مواردی که حتماً باید در Theme باشد: Visual Defaults - Background - Font - Colors (JSON Theme)

3.5.1 در این مرحله، ساختار کلی داشبورد با هدف هدایت تدریجی کاربر از شاخص‌های کلیدی به تحلیل‌های جزئی طراحی شد. چیدمان اجزا بر اساس اولویت تصمیم‌گیری و الگوی حرکت چشم کاربر تعریف گردید تا بیشترین وضوح و کمترین بار شناختی ایجاد شود.

3.6.1 این مرحله، صفحات داشبورد بر اساس سؤالات تجاری طراحی و برای هر سوال، مناسب‌ترین نوع نمودار انتخاب شد. انتخاب Visual ها با هدف بیشینه‌سازی وضوح، کاهش بار شناختی و پشتیبانی مستقیم از تصمیم‌گیری انجام گرفت.

3.7.1 و در انتها با توجه به تحقیقات و داده های بدست آمده و جلسات با کارفرما یک گزارش کامل و اصولی برای ردمی گیت امده کرده که ساختار گزارش تا این قسمت را شرح دهد و حتما برای داده بیشتر به این قسمت هم ارجاع دهید و در نهایت کامیت بخش سوم

4. (Implementation)پیاده‌سازی گزارش

4.1.1 ضبط ویدیو از طول فرایند اجرایی در طول اجرای این مرحله، فرآیند کامل مهندسی داده و مدل‌سازی به صورت ویدیو ضبط شد. هدف از ضبط، مستندسازی و ارائه شفاف فرآیند برای انتشار آموزشی در فضای مجازی بود. برای ضبط ویدیو از نرم‌افزار Techsmith استفاده شده است. لازم به ذکر است که محتوای ضبط شده تنها شامل مراحل اجرایی و تغییرات داده‌ها است و هیچ اطلاعات حساس یا خصوصی منتشر نشده است.

4.2.1 در این مرحله، اجزای مشترک صفحات شامل هدرها، فیلترها، بوکمارک‌ها و الگوهای تعاملی به‌صورت یکپارچه پیاده‌سازی شد تا انسجام بصری، سهولت نگهداری و سرعت توسعه داشبورد افزایش یابد.

4.3.1 ر این مرحله، هر صفحه داشبورد به‌صورت مستقل و بر اساس سؤالات تجاری پیاده‌سازی شد. پس از تکمیل هر صفحه و اطمینان از صحت محاسبات، کیفیت بصری و تعاملات، وضعیت پروژه در مخزن Git ثبت گردید.

4.4.1 ر این مرحله، بازنگری فنی اولیه با تمرکز بر رنگ‌بندی، تایپوگرافی، اندازه و چیدمان المان‌ها انجام شد تا خوانایی و انسجام بصری داشبورد بهبود یابد.

4.5.1 در این مرحله، نمودارها و جریان داستان داده بازبینی شد تا روایت داشبورد شفاف‌تر شده و پیام‌ها به‌صورت مرحله‌به‌مرحله و تصمیم‌محور به مخاطب منتقل شوند.

4.6.1 در انتها با توجه به تحقیقات و داده های بدست آمده و جلسات با کارفرما یک گزارش کامل و اصولی برای ردمی گیت امده کرده که ساختار گزارش تا این قسمت را شرح دهد و حتما برای داده بیشتر به این قسمت هم ارجاع دهید و در نهایت کامیت بخش چهارم

5. معرفی و ارائه پروژه

5.1.1 تهیه اسلاید معرفی پروژه آماده‌سازی یک اسلاید متنی برای معرفی کلی پروژه، هدف آن و خروجی‌ها

5.2.1 تولید محتوای معرفی پروژه تهیه یک پست شامل: توضیح فرایند انجام پروژه ویرایش محتوای بصری نوشتن متن توضیحی (Caption) برای انتشار

5.3.1 ضبط ویدیو معرفی به زبان انگلیسی ضبط یک ویدیو به زبان انگلیسی شامل ارائه پروژه، توضیح رویکرد کاری و دلایل انتخاب هر بخش.