

🚀 عنوان پروژه : واکشی دیتا در Power BI و ETL

کردن آن

• تاریخ شروع: Dec 4, 2025

• لینک گیت‌هاب: [لینک](#)

• لینک ویدیو/ارائه: هنوز کامل نیست

🚀 فاز ۰: پیش‌زمینه استراتژیک

۱. ☒ ساخت تمپلیت و ساختار پوشه

- powerbi-etl-orders/
 - |— data/ # تقویم, List of order, order breakdown, فایل‌های اکسل خام شمسی
 - |— notebooks/ # (اختیاری Python اگر برای تحلیل جانبی از)
 - |— scripts/ # جانبی داشته باشیم Python یا M اگر اسکریپت
 - |— pbix/ # پروژه.pbix (فایل‌های)
 - |— docs/ # ... و notes.md, troubleshooting.md
 - |— outputs/ # ها export، اسکرین‌شات داشبورد، نمودارها
 - |— README.md
 - |— requirements.txt # داری Power BI فراتر از Python/M فقط اگر کار

۲. ☒ تعریف سبک شخصی (Personal Signature)

- برندینگ نوت‌بوک: (مثال: لوگو یا بلوک Markdown ثابت در ابتدای نوت‌بوک)
- برندینگ بصری: (مثال: تم رنگی ثابت برای نمودارها یا قالب Canva برای پست‌ها)
- داخل Power BI: یک تم رنگی ثابت برای ویژوال‌ها (Theme) + لوگو کوچک در گوشه داشبورد.
- در نوت‌بوک‌ها/داک‌ها: یک بلوک Markdown ثابت با عنوان پروژه، نام خودت، تاریخ و لینک گیت‌هاب.

🔍 فاز ۱: قبل از انجام تسک (تحلیل و آماده‌سازی)

۱. ✓ تعریف هدف کسب‌وکار (Business Question)

- سوال اصلی که باید پاسخ داده شود چیست؟
 - دیتای سفارش‌ها را طوری ETL کنم که بتوانم به سادگی سود (Profit) را بر اساس کشور، محصول و زمان تحلیل کنم، محصولات/کشورهای با سود منفی را شناسایی کنم و امکان گزارش‌دهی شمس‌ی برای کاربر نهایی وجود داشته باشد.

۲. ✓ شناسایی معیارهای موفقیت (KPIs)

- موفقیت این تحلیل چگونه اندازه‌گیری می‌شود؟
 - مجموع Sales و Profit
 - درصد سفارش‌های با Negative profit
 - سود به ازای هر کشور
 - سود به ازای هر محصول
 - روند زمانی فروش/سود (بر اساس تاریخ شمس‌ی بعد از تبدیل)

۳. ✓ پیش‌بینی ساختار و نیاز داده‌ها

- چه داده‌هایی نیاز است؟ کیفیت مورد انتظار چیست؟
 - جداول: List of order, orderbreakdown, فایل تقویم شمس‌ی
 - ستون‌های حیاتی: Order Date, Country, Product name, Sales, Profit, ...
 - کیفیت مورد انتظار: بدون NULL در ستون‌های کلیدی، نوع داده درست (تاریخ، عدد، متن)، فرمت تاریخ استاندارد.

۴. ✓ انتخاب استراتژی و مدل

- از چه روشی استفاده خواهد شد؟
 - ETL کامل در Power Query Editor
 - پاک‌سازی و استانداردسازی (Detect Data Type, Replace, Split)
 - ساخت ستون جدید Profit Category
 - Append و Merge (برای ساخت جدول Consolidated Total)
 - افزودن تقویم شمس‌ی و ساخت Relationship

○ فعلاً مدل تحلیلی ساده (Aggregations و Slicers) نه مدل آماری پیچیده.

۵. بررسی پروژه‌های مشابه (Benchmarking) ☒

- چه درس‌هایی از پروژه‌های مشابه گرفته شد؟ (لینک‌های مفید)

1. پروژه اول: داشبورد فروش کافی‌شاپ (Coffee Shop Sales) (Dashboard)

منبع/مدرس: کانال Data Analysis Tech در یوتیوب

لینک: Power BI برای مبتدیان: تحلیل فروش کافی‌شاپ

خلاصه پروژه: این پروژه بر داده‌های واقعی فروش یک کافی‌شاپ زنجیره‌ای (Maven Roasters) در سه شعبه طی ۶ ماه (ژانویه تا ژوئن ۲۰۲۳) متمرکز است medium.com. در این آموزش ویدئویی نحوه استخراج داده‌های تراکنشی از فایل Excel/دیتابیس، پاکسازی و تبدیل آن‌ها با Power Query Editor و سپس ساخت داشبورد تعاملی در Power BI نمایش داده می‌شود. برای تحلیل زمانی، یک جدول تقویم سفارشی (Date Table) در Power BI ایجاد و با جدول تراکنش‌ها مرتبط شده است تا امکان فیلتر گذاری و محاسبات زمانی فراهم شود github.com. در طی فرایند Transform، ستون‌ها نام‌گذاری مجدد و نوع داده آن‌ها تنظیم شده (مثلاً تبدیل متن به عدد و تاریخ) و رکوردهای ناقص یا تکراری بررسی و رفع شده‌اند github.com. نتیجه نهایی یک داشبورد پویا با شاخص‌های کلیدی عملکرد (KPI) مانند مجموع فروش، تعداد سفارش‌ها، الگوی فروش به تفکیک روزهای هفته/آخر هفته و پرفروش‌ترین محصولات است. این داشبورد به مدیر فروش امکان می‌دهد روندها و الگوهای مهم را شناسایی کند medium.com.

2. پروژه دوم: داشبورد تحلیل منابع انسانی (HR Analytics) (Dashboard)

منبع/مدرس: کانال Codebasics در یوتیوب (پروژه چالشی Pinali Mandalia)

لینک: پروژه تحلیل منابع انسانی با Power BI

خلاصه پروژه: این پروژه یک داشبورد جامع منابع انسانی را پیاده‌سازی می‌کند که به صورت end-to-end فرایند ETL و گزارش‌سازی را پوشش می‌دهد. داده‌های واقعی HR شامل جزئیات پرسنل، حضور و غیاب، بازخوردها و ترجیحات کاری کارمندان مورد استفاده قرار گرفته است

github.com. هدف پروژه، خودکارسازی گزارش‌های کلیدی واحد منابع انسانی (مانند نرخ مرخصی استعلاجی، ترجیحات دورکاری/حضور و دلایل آن) بوده است. در بخش تبدیل داده‌ها با Power Query، ابتدا فایل‌های Excel خام (شامل شیت‌های حضور و غیاب سال‌های 2022-2023 و غیره) پاکسازی شده‌اند: رکوردهای تکراری و مقادیر Null حذف گردیده و فرمت‌ها استاندارد شدند. سپس با Merge Queries اطلاعات پراکنده در چند جدول ترکیب شده است؛ برای مثال، جدول حضور و غیاب با جدول اطلاعات کارمندان ادغام شده تا فیلدهای تکمیلی (مانند دپارتمان یا موقعیت شغلی) به یک جدول واحد اضافه شود github.com. همچنین چند ستون سفارشی (Conditional Column) ایجاد شده است (مانند ستونی که دسته‌بندی انواع مرخصی یا دلایل دورکاری را مشخص می‌کند) github.com. پس از Transform، مدل داده به شکل اسکیمای ستاره‌ای طراحی شده است: جداول بعد (بخش‌ها، کارمندان، محل خدمت) به جدول حقیقت حضور و غیاب و بازخورد متصل شده‌اند تا روابط یک‌به‌چند به‌درستی تعریف شوند github.com. در نهایت در محیط گزارش‌سازی، معیارهای مهم با DAX محاسبه و کارت‌های KPI (مثل درصد مرخصی استعلاجی، درصد دورکاری) و نمودارهای تعاملی (مانند مقایسه روند ماهانه دورکاری در هر دپارتمان) ساخته شده است. این داشبورد نهایی به تیم HR امکان می‌دهد به‌جای تهیه دستی گزارش‌ها، به‌صورت لحظه‌ای وضعیت نیروی کار را رصد و تصمیمات مبتنی بر داده بگیرند github.com.

۶. مستندسازی اولیه (در notes.md یا این داک) ☒

- فرضیات (Assumptions): داده‌ها کامل هستند یا اگر ناقص‌اند، ردیف‌های ناقص حذف می‌شوند. واحد پولی همه رکوردها یکسان است.
- ریسک‌ها (Risks): اشتباه در تبدیل تاریخ میلادی-شمسی اشتباه در Merge (Join) نوع اشتباه و تکراری شدن ردیف‌ها
- سؤالات باز (Open Questions): آیا کاربر نهایی فقط نیاز به گزارش سطح بالا دارد یا تحلیلی جزئی‌تر هم لازم است؟

فاز ۲: حین انجام تسک (اجرا و مستندسازی)

- ۱. ☒ مدیریت و ردیابی پیشرفت
- ☒ ابزار انتخابی: Trello
- ☒ تسک‌های اصلی خرد شده: در ترلود خورده شده به 19 قسمت که هر کدام زیر تسک‌های خودش را دارد

☐ ۲. اجرای کد تمیز و داستانی

-
-

☐ ۳. کامنت‌گذاری بر تصمیمات (Rationale)

-
-

☐ ۴. ثبت مشکلات و راه‌حل‌ها (در `troubleshooting.md`)

• ثبت مشکلات و راه‌حل‌ها

○

☐ ۵. نسخه‌گذاری هوشمند در گیت‌هاب (Smart Commits)

۱. برای مدیریت نسخه‌ها، پس از هر مرحله مهم یک Commit معنادار در گیت‌هاب ثبت شده است:

۲. Initial commit: ایجاد ساختار پروژه و افزودن فایل‌های گزارش اولیه.

۳.

☐ ۶. ساخت داشبورد (اختیاری/ارزشمند)

-

☐ ۷. مدیریت محیط (requirements.txt)

• (تیک بزنید که فایل نیازمندی‌ها ایجاد/به‌روزرسانی شده است.)

📢 فاز ۳: بعد از انجام تسک (ارائه و شبکه‌سازی)

۱. به‌روزرسانی گیت‌هاب با `README.md` قوی

☐ چک‌لیست README:

☐ مشکل (Problem)

☐ متدولوژی (Methodology)

☐ یافته‌های کلیدی (Key Findings)

☐ چالش‌ها و راه‌حل‌ها (از `troubleshooting.md`)

☐ توصیه‌های کسب‌وکاری (Actionable Recommendations)

۲. تولید ویدیوی ارائه فنی (فرمول ۵ بخشی)

☐ چک‌لیست ویدیو:

- ☐ معرفی (۳۰ ثانیه): من کی‌ام و پروژه چیست؟
- ☐ هدف کسب‌وکار: (چرا این مهم بود؟)
- ☐ فرایند و چالش‌ها: (داستان مشکلات و حل آن‌ها)
- ☐ نتایج عددی و بصری: (مهم‌ترین نمودارها و اعداد)
- ☐ درس‌های آموخته شده: (نتیجه‌گیری)
- ☐ لینک ویدیو: لینک

۳. پست‌گذاری استراتژیک در لینکدین (Storytelling)

- پیش‌نویس پست (تمرکز بر داستان، نه گزارش):
 - [پیش‌نویس شما... مثال: "در این پروژه با چالش X روبرو شدم، اما با روش Y توانستم الگوی Z را پیدا کنم..."]
- لینک پست منتشر شده: [لینک]

۴. دریافت بازخورد و ریکامندیشن (منتورینگ دوطرفه)

- ☐ مرحله ۱: ارسال پروژه برای منتور و درخواست بازخورد.
- ☐ مرحله ۲: اعمال بازخوردها و نشان دادن نتیجه به منتور.
- ☐ مرحله ۳: درخواست ریکامندیشن نهایی بر اساس کار بهبودیافته.

🧠 فاز ۴: تحلیل بعد از پروژه (یادگیری شخصی)

- (این بخش برای ژورنال شخصی شماست و نیازی به انتشار ندارد)

۱. بازبینی شخصی (Self-Retrospective)

- چه چیزی را بهتر می‌توانستم انجام دهم؟
 - [یادداشت شما...]

- کدام مهارت جدید را یاد گرفتیم یا تقویت کردیم؟
 - [یادداشت شما...]
- کدام بخش را باید در پروژه بعدی بهبود بدهم؟ (مثال: اتوماسیون بیشتر، تصویرسازی بهتر)
 - [یادداشت شما...]