

مقدمه:

- جزء خانواده ETL هاست:
ETL : Extract Transform Load (Power Query , SSIS)
بخواند
- منابع مختلف : اکسل / وب / دیتابیس
سر و سامان / تمیز / cleaning
لود/ مقصد:
- همان اکسل است.
- اکسل: Pivot Table / Power Pivot

Power Query (ETL) : Excel , Power BI, ...

چرا و کی؟

همیشه هر چیزی با pq مگر آنکه خلافتش ثابت بشه.
تقریباً : امکانات اکسل کار با دیتا را داشته باشد :

remove duplicate, text 2 columns, value , find replace,
vlookup, pivot ,

تقریباً برخی دستورات اصلی sql هم دارد:

اکسل:

فرمت سلول / چارت / pdf و

تاریخچه : ۲۰۱۳ اومد با نصب شود / سپس از ۲۰۱۶ به خود اکسل اضافه شد / ۲۰۱۰ هم می شود نصبش کرد/

وقتی که دیتا خیلی زیاد است

تعداد سطر ۱ میلیون است .

سرعت بسیار بالایی دارد.

وقتی استفاده می کنیم که می خواهیم کاری روی دیتا را اتوماتیک کنیم:

برنامه نویسی می کنیم / زبان برنامه نویسی است به نام Mashup M:

محدودیت دارد:

مرتفع می کند.

میلاد عسگری : میلاد عسگری : fuzzy match

کپی پیست زیر هم .

که zero code

مشکل:

نمی شناسند.

کاربرد

شروع کار:

پاورکوئری دیتا می خواهند :

- می تواند دیتا را از **همین فایل اکسل** که الان توی RAM بخواند.
- حتما باید Table شده باشد.
- یا named range باشد.
- شیت نمی تواند بخوان!
- فایل
- دیتایی که روی حافظه جانبی ذخیره شده است. xls, txt, pdf, xlm
- می تواند شیت یک فایل دیگر را بخواند.
- دیتابیس
- عجیبی
- نرم افزاری بنویسید

farsaran.com/link : northwind 2000

مسئله) برنامه ای بنویسید که روی دیتا ۵ آیتم پرفروش فصل ۴ را نمایش دهد.
فیلتر : معادل حذف / delete /
پاورکوئری به حروف A / a حساس است.

edit: alt + F12

یافتن منشاء خطا (اسم تیبل / اسم ستون تغییر می کند)

بدیهی است که وقتی کد می نویسیم / اتوماتیک کنیم که حتما نظمی دارد.

کوئری چیست؟

برنامه ای که روی دیتا کاری را انجام می دهد.

حاصل کوئری فرمول اکسلی بنویسیم / Calculated column

آشنایی با transformation های ساده

تعیین نوع دیتای هر ستون

شرترین تبدیل همین changed type

چرا: چون جنس ستونها را حدس می زد ۲۰۰ سطر

بسیار cpu را اشغال می کند

الکی باشه / کم بشه : وقتی انجام می دهیم که :

آخر سر / حذف / تمیز /

وقتی مجبوریم : vlookup

تعیین نوع ستونها الزامی است و گرنه محاسبات در اکسل انجام نمی شود و همچنین برخی از کارها مثل جمع زدن در پاورکوئری انجام نمی شود.

Query Options -->

Never detect column types and headers for unstructured sources

تعیین تایپ تاریخ میلادی / شمسی

🔗 توجه مهم: این ربطی به تنظیمات Control Panel دارد.

