

The Analyst Process: Converting a Past Report to Power BI

www.plaid.is

Without a Centralized Metadata Management Tool



The “Scavenger Hunt” for Data

Spreadsheets with in-cell formulas

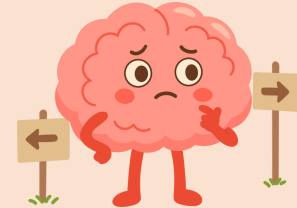
Ad-hoc SQL queries

Locally stored data files

An uncommented Python script

Inconsistently named and dated versions of dashboard

Email threads



Missing, outdated, or conflicting documentation

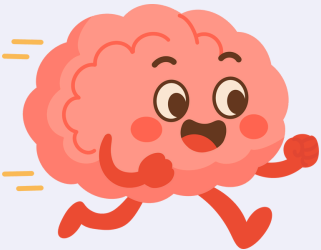
With a Centralized Metadata Management Tool

With Plaid Govern

A Centralized Metadata Management tool

With Plaid Govern & Certified Datasets built for Power BI

(see Plaid Data Integration & Warehousing)



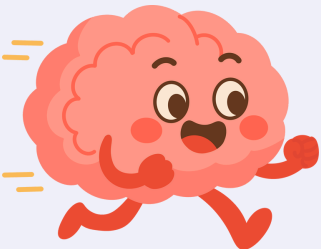
Refer to existing data dictionary for trusted KPI definitions, tables, and locate necessary definitions, data sources, and fields.

Manual data prep and cleaning in Power Query to create trusted dataset

Connect to dataset



Interpret multiple script languages for calculations and transforms



Metadata dictionaries provide clarity in non-coded language with standard naming conventions for simplified identification

Recreate logic using Power Query functions and M scripts

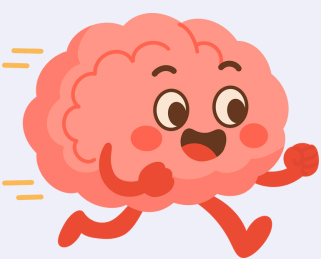
Build data model



Analyze original data source structure.

Identify underlying data relationships by manually combing through database tables for primary and foreign keys.

Translate to a star schema.



Refer to existing lineage map and data dictionary.

Identify additional compound keys, sort tables, or bridge tables Power BI requires

Build star schema

Write DAX measures

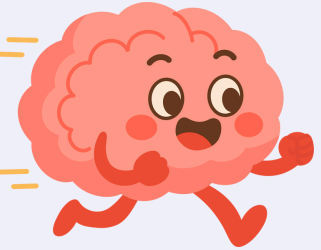


Validate and determine correct KPI metric calculation in original report data set.

Translate from other query language to DAX

Go back and forth between different reports and data model build to ensure row and filter contexts are applied appropriately.

Discover inconsistent or illogical calculations in original. Go back to manual data prep.



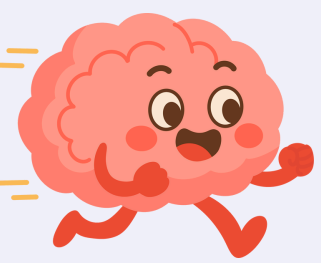
Refer to data dictionary for trusted metric definitions. Use lineage map to quickly trace row and filter contexts.

Build visuals

Publish Workbook



Explain data and metrics to end user. Repeat in full every cycle.



End users use data dictionary and lineage maps to understand dashboard.

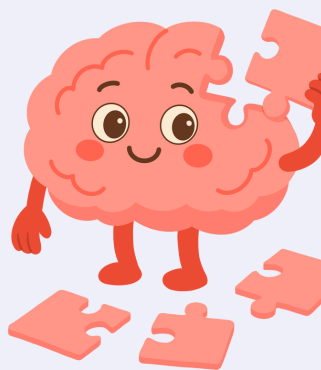
(in decreasing amounts as time goes on and trust builds in reporting)

Create Process Documentation



Update documentation manually.

Often rushed and incomplete, or skipped altogether, resulting in a future scavenger hunt. Usually not accessible to end users.



Use Plaid Govern to automatically ingest new structure.

Add new definitions to dictionary only if new fields created.

