Name:

Company Mission:

FoxTeam wishes to enter the E-sports industry but is having trouble with their data handling methods. They need a system that supports secure warehousing of E-sports related data. They expect to be able to easily access the database. Furthermore, they expect the system to build reports that are consistent and topical.

Manual cleaning:

\u00e2\u20ac\u2122

\*x

“w|w|w|w”

“|”

Boolean Column, the Multiple Column, or the Bridge Table method

Needs

Enron

Trending analysis

-Data source

-PandaScore

-OLAP

-ETL tool

-data reporting (Spelunk)

-charts, tables, etc

-data from same time

-drill down

-atomize data

Wants

-dummy data from hltv (CS go)

-PandaScore

-OLTP

-spelunk

Data Flow Diagram

Retrieval

GitHub (Raw\_GitHub\_Datatset.xlsx) > CSV (Raw\_GitHub\_Datatset.csv) >

Filtering:

* unwanted fields

> Remove\_column.py > buffer.csv

1st Transformation:

* csv to json

>CSV\_to\_JSON.py > dnd\_players.json

Cleaning/formatting:

* unwanted characters
* separates into somewhat normalized tables in JSON (delineated by newlines)
  + dimension
  + fact
  + bridge

>Clean\_and\_Make\_KVPairs > various tables in json

Load:

Manually uploading into BigQuery

Data Reporting:

Data Studio manually exporting from BigQuery to DataStudio