



Photograph by Markus Spiske via Unsplash.
(https://unsplash.com/photos/text-50fY_4ckUw).

Graphs that Speak

Methodology

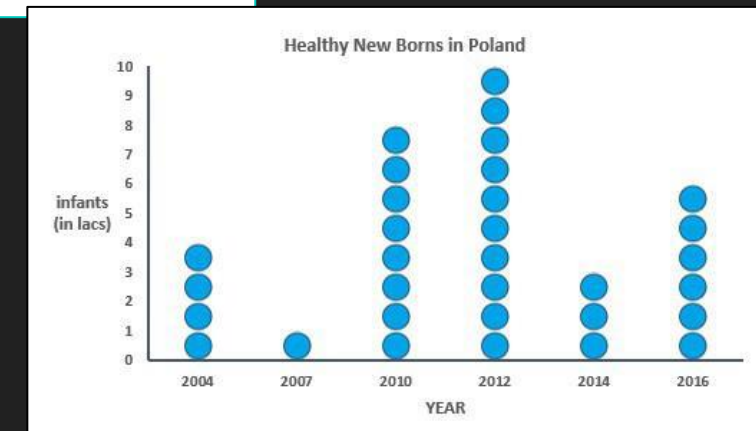
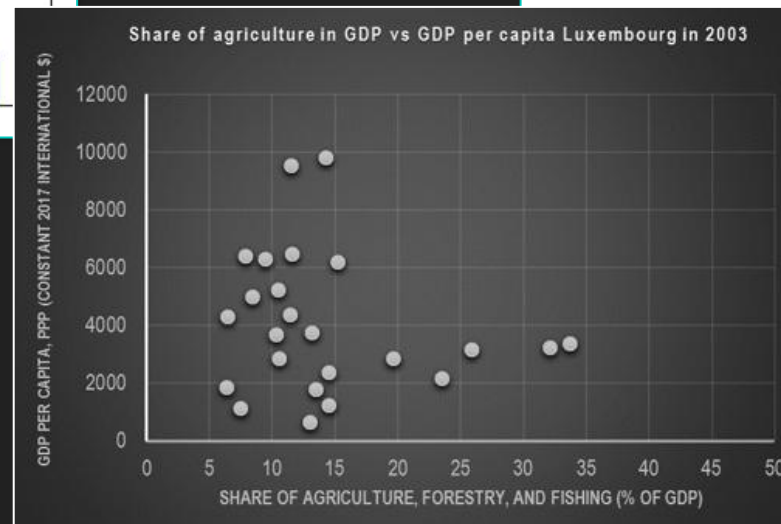
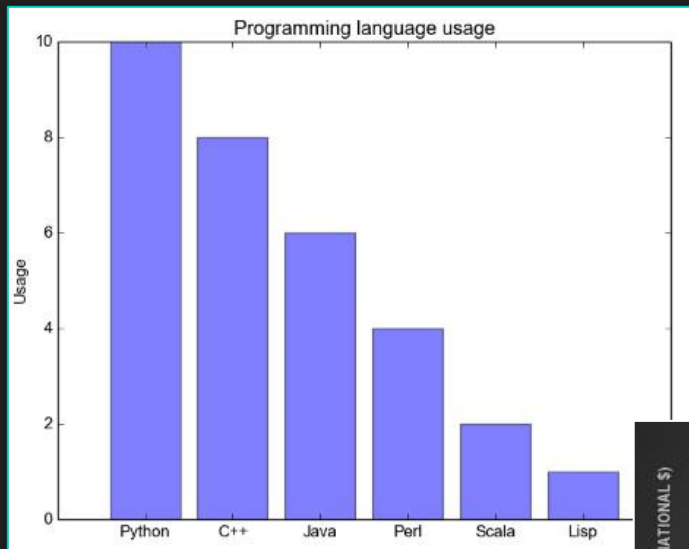
Outline

- 1) Extracting the Annotations
- 2) Prioritized Data Elements
- 3) Model Implementation

Methodology

- GOAL = successfully interpret the data series depicted in the test set
- Dataset
 - 60,578 synthetic and professionally produced .jpg images
 - 60,578 associated annotations in .json format

Examples of Charts



1) Extracting Annotations

- One .json file

- Source :

- Generated
 - Extracted

- Chart Type :

- Horizontal Bar
 - Vertical Bar
 - Scatter
 - Dot

- Data Series 'X' :

- x values of the graph

- Data Series 'Y' :

- y values of the graph

VERTEX AI

EXTRACT

Data Frame

Data Frame							

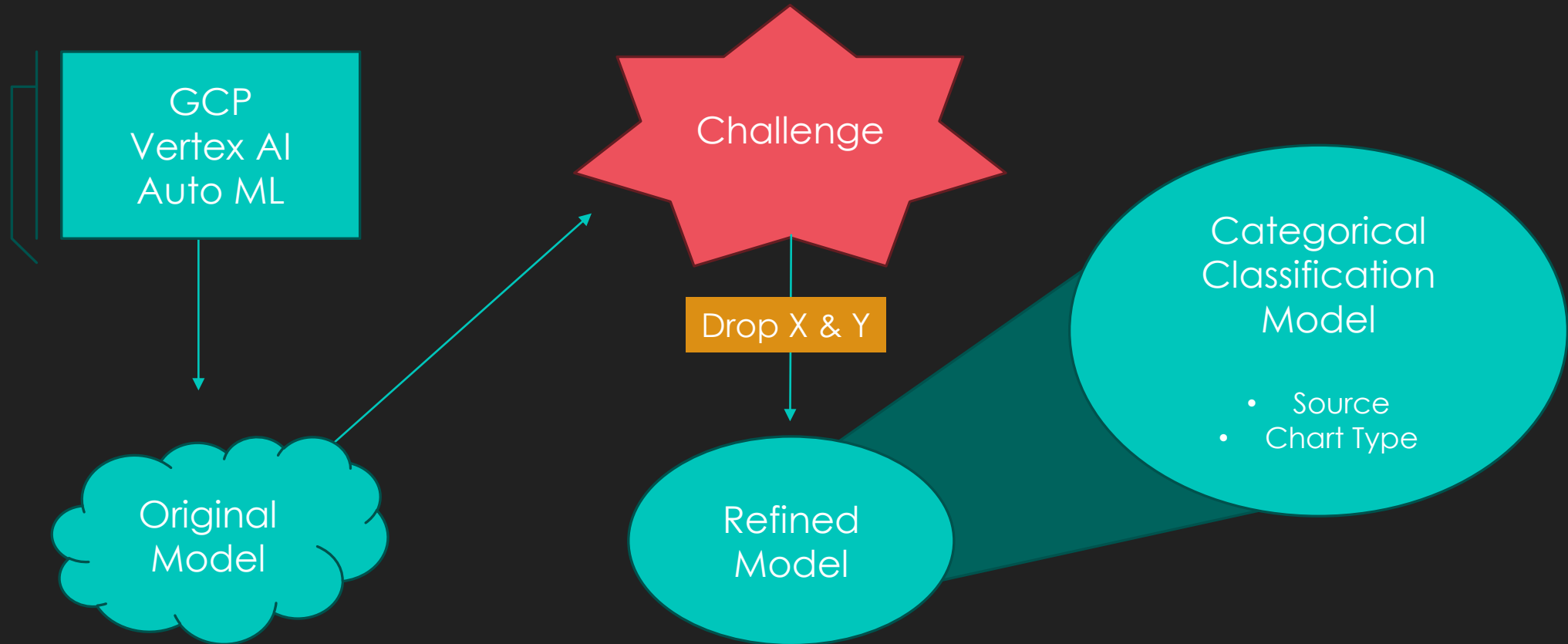
.CSV

.CSV							

2) Prioritized Data Elements

- Draft code in Python notebook
- Transfer images and .csv to GCP Storage Bucket
- Conduct Exploratory Data Analysis
- Build and deploy model with Vertex AI AutoML tool

3) Implementation



Exploratory Data Analysis

Distribution of Chart Type

