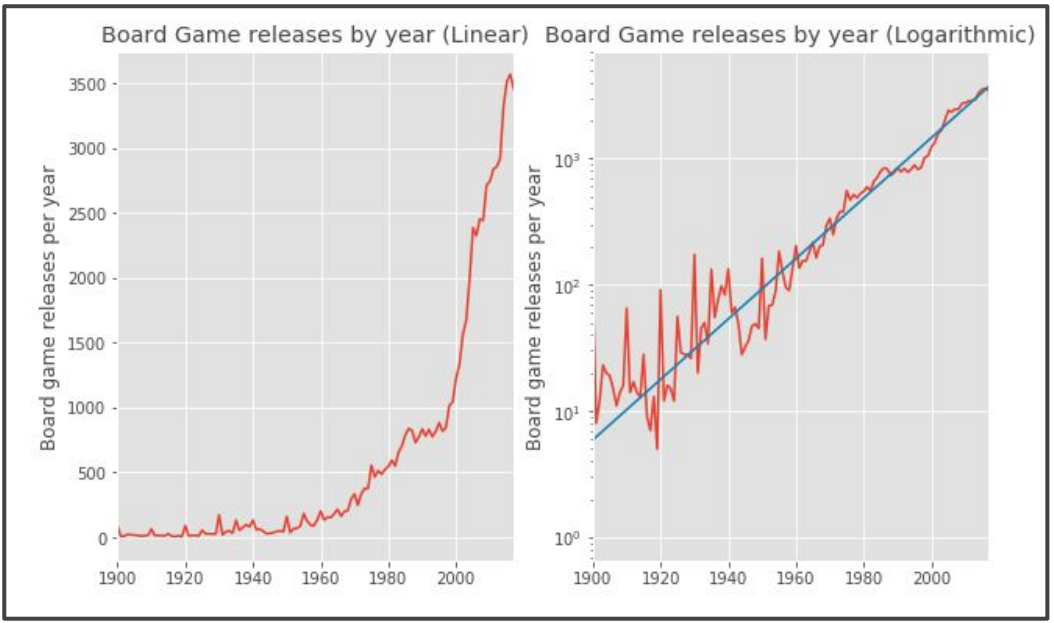
BoardGameGuru

“The Guru to your Geek”

The Rise of Boardgames



The Number of Boardgames is growing exponentially.

In 2015, there were ~3,500 Boardgames published. That's *ten a day*.

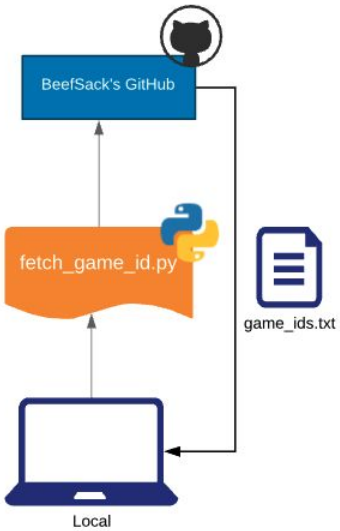
Graph Credit: Dinesh Vatvani

Image Credit

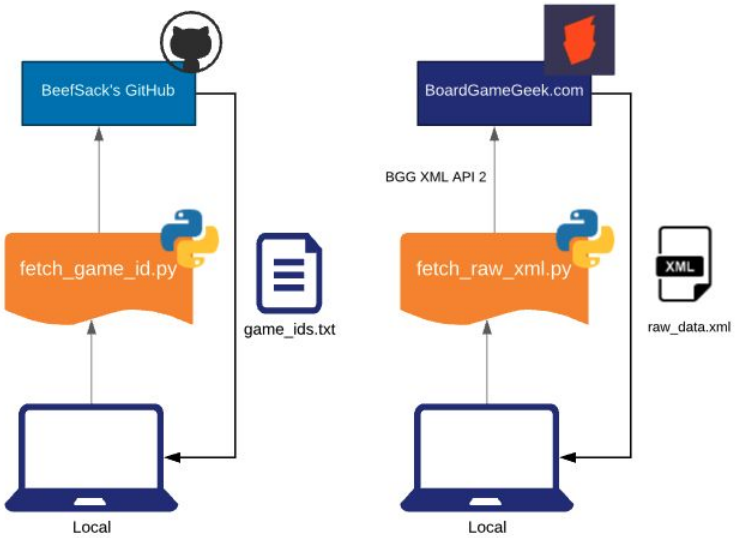
BoardGameGuru = A Boardgame Recommendation Engine

Demo

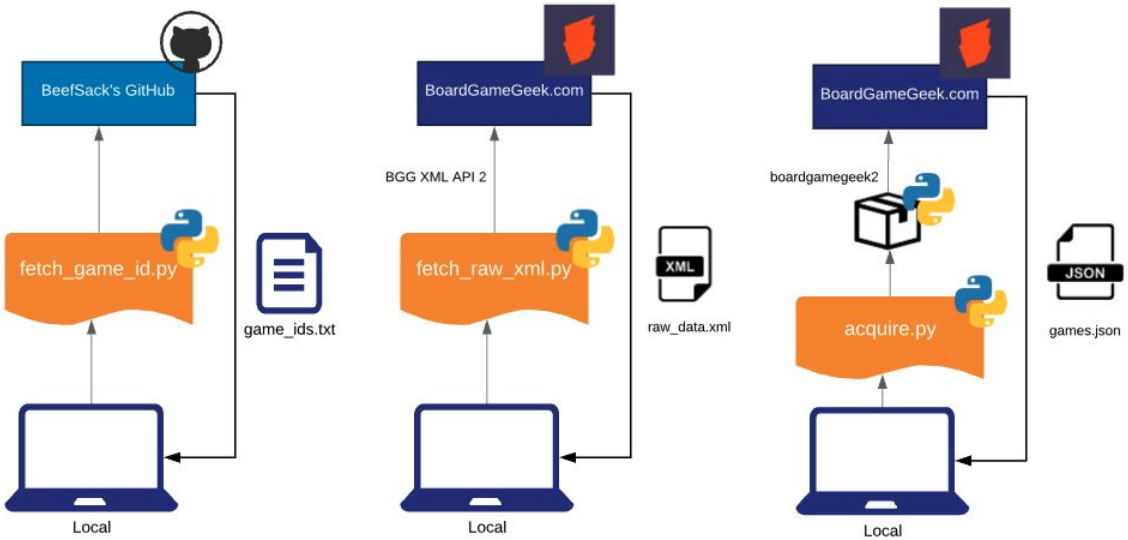
Data, Workflow & Architecture

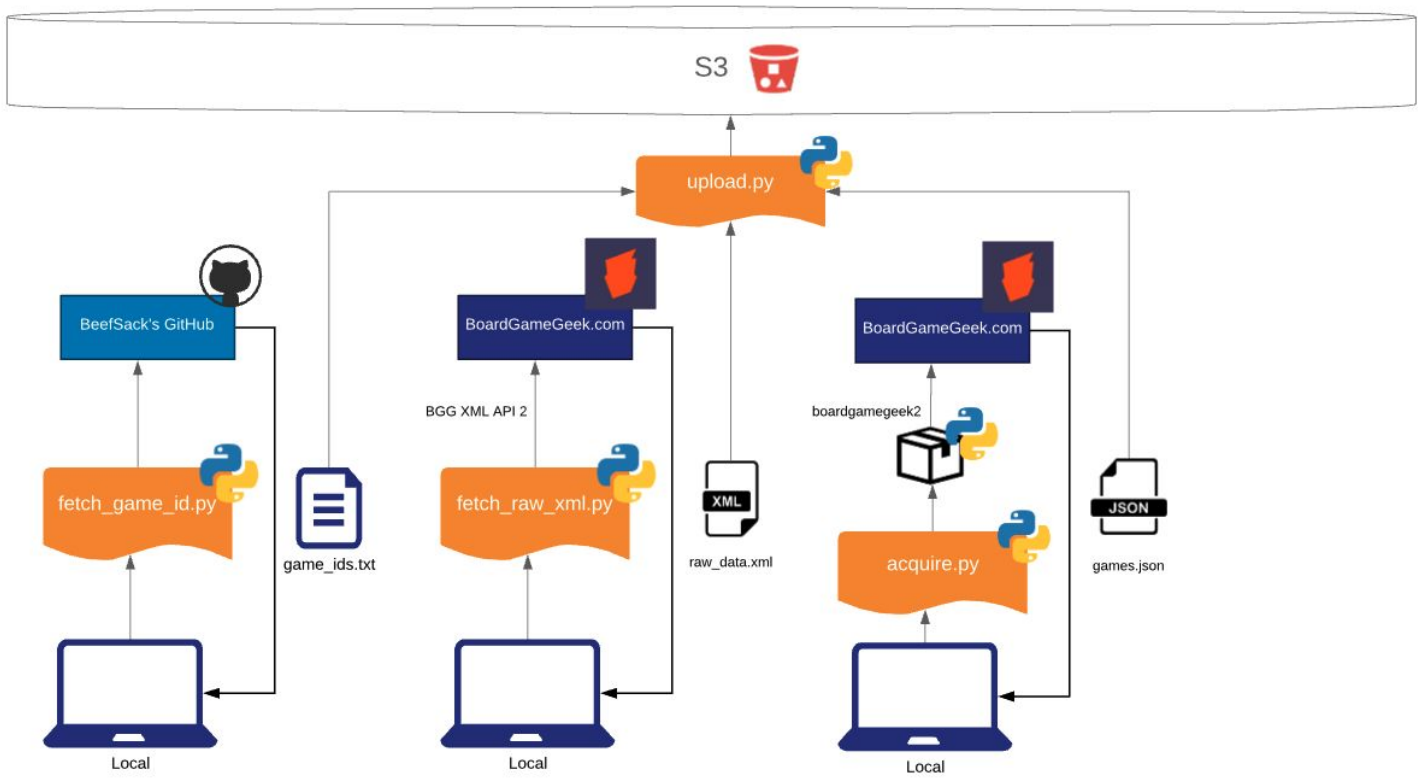


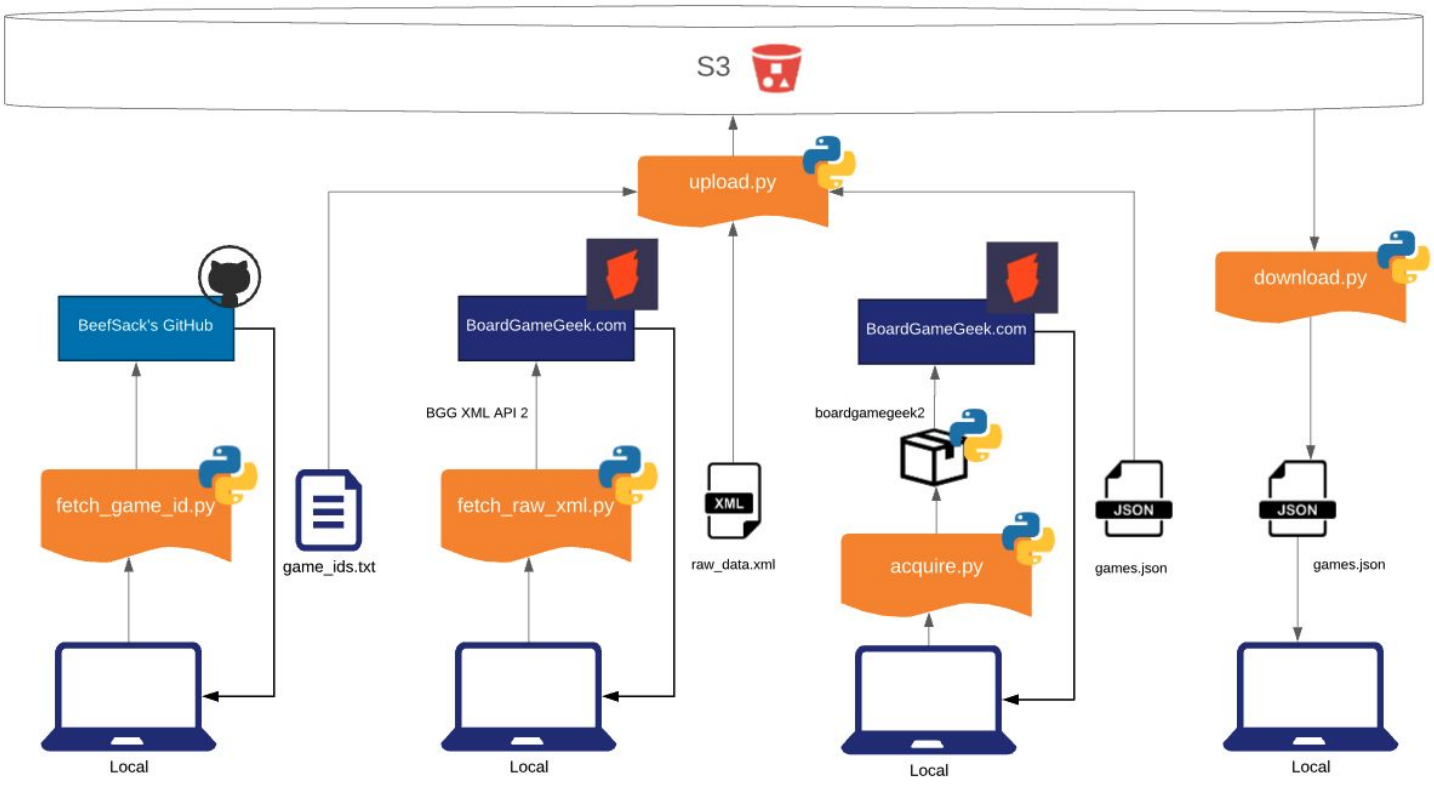
Data, Workflow & Architecture

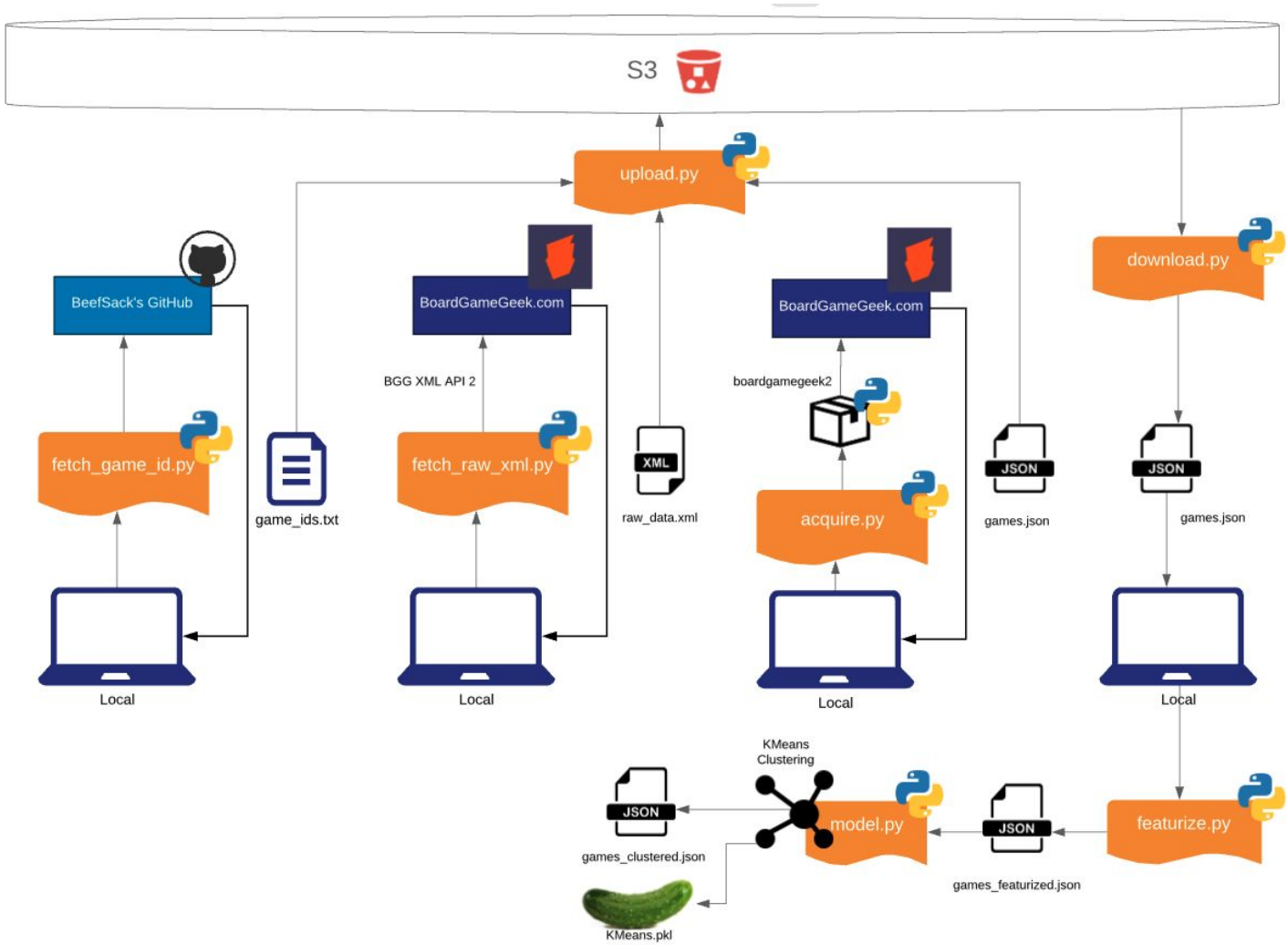


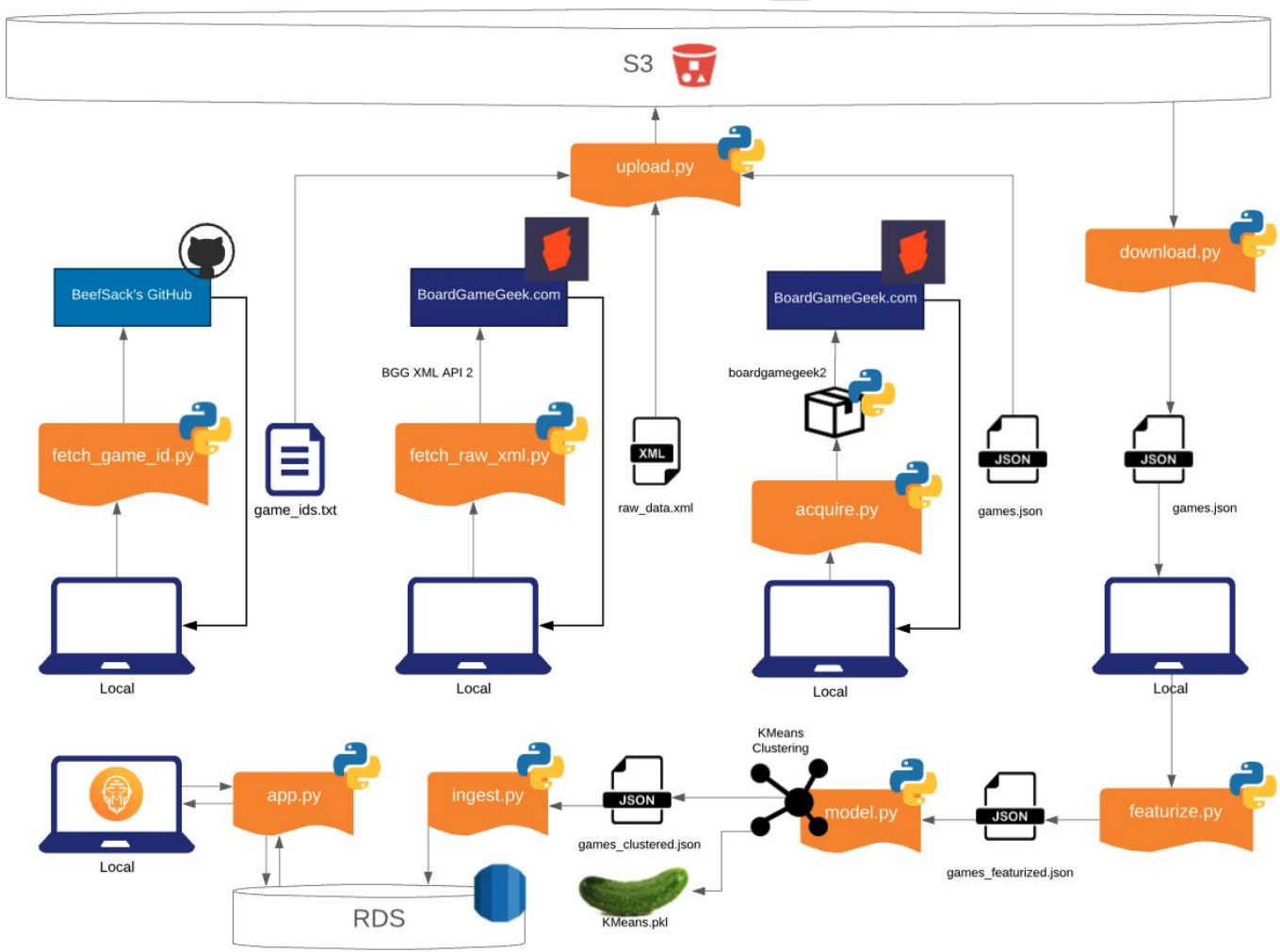
Data, Workflow & Architecture

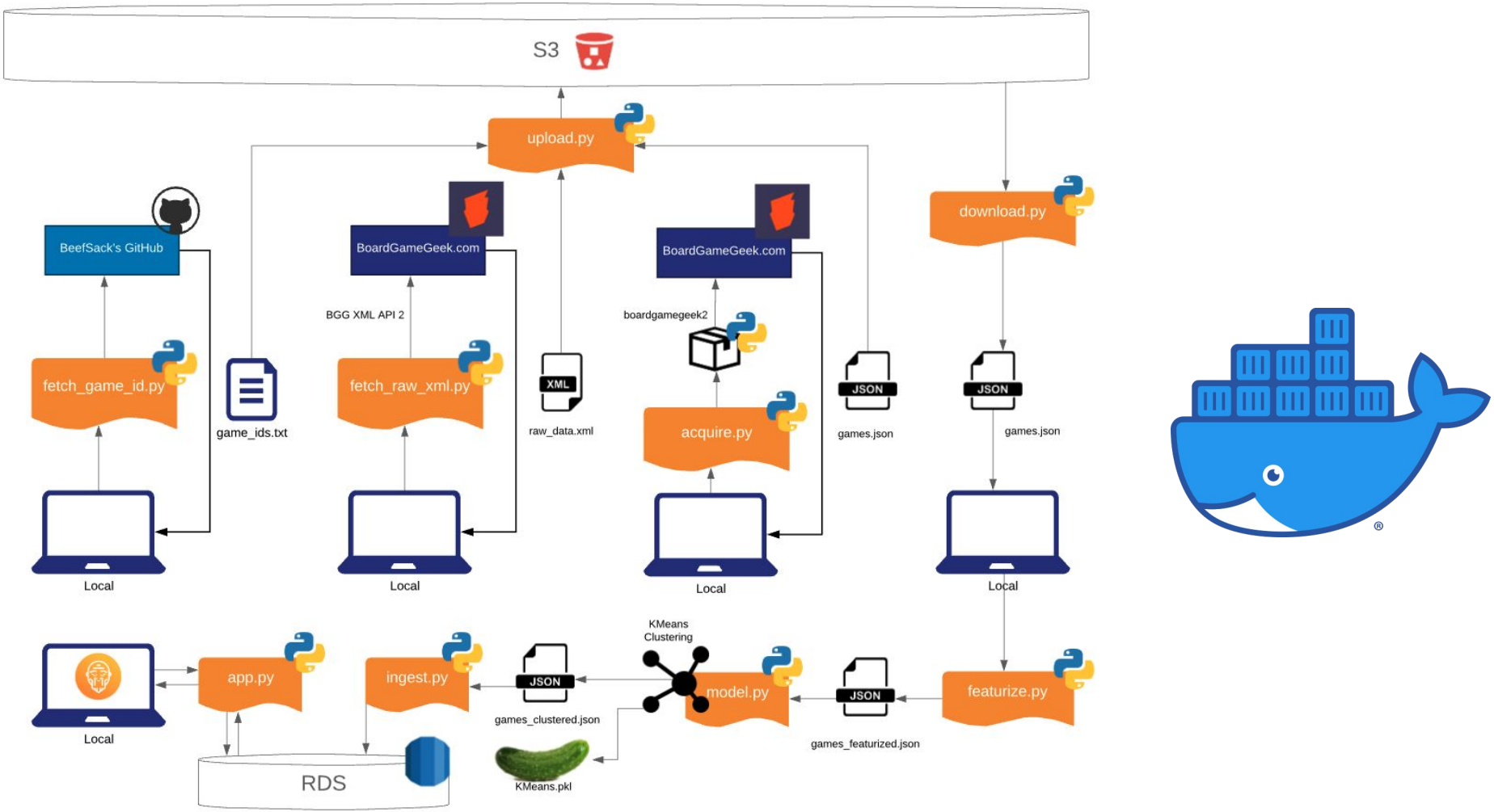



S3
upload.py
BeefSack's GitHub
BoardGameGeek.com
BoardGameGeek.com
BGG XML API 2
boardgamegeek2
fetch_game_id.py
game_ids.txt
fetch_raw_xml.py
XML
raw_data.xml
acquire.py
JSON
games.json
Local
Local
Local

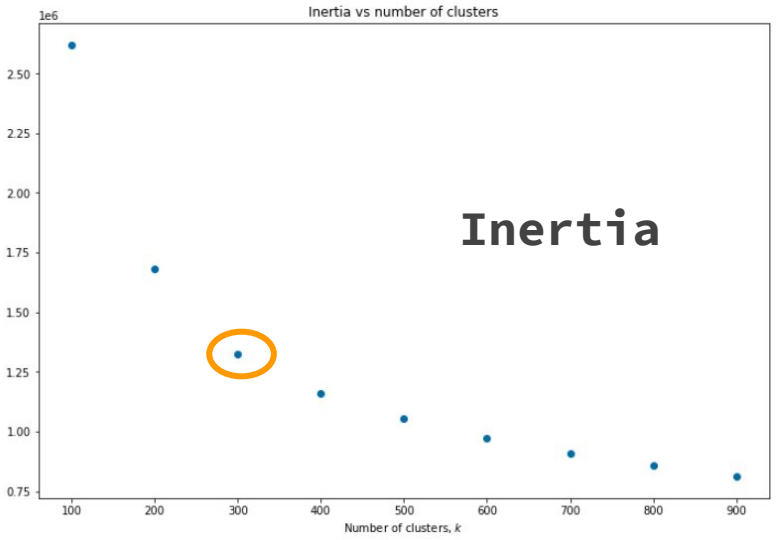

S3
upload.py
download.py
BeefSack's GitHub
BoardGameGeek.com
BoardGameGeek.com
BGG XML API 2
boardgamegeek2
fetch_game_id.py
game_ids.txt
fetch_raw_xml.py
raw_data.xml
acquire.py
games.json
games.json
Local
Local
Local
Local

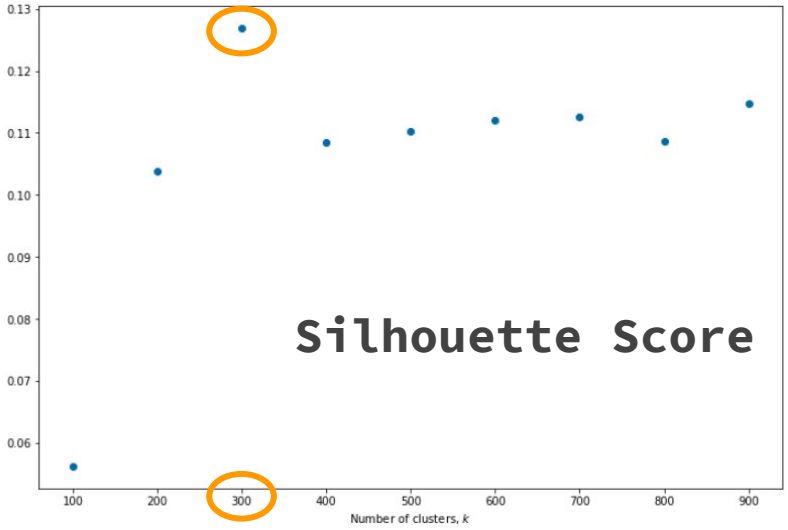

S3
upload.py
download.py
BeefSack's GitHub
BoardGameGeek.com
BoardGameGeek.com
BGG XML API 2
boardgamegeek2
fetch_game_id.py
game_ids.txt
fetch_raw_xml.py
XML
raw_data.xml
acquire.py
JSON
games.json
JSON
games.json
Local
Local
Local
Local
KMeans
Clustering
JSON
games_clustered.json
model.py
JSON
games_featurized.json
featurize.py
KMeans.pkl


S3
upload.py
download.py
BeefSack's GitHub
BoardGameGeek.com
BoardGameGeek.com
BGG XML API 2
boardgamegeek2
fetch_game_id.py
game_ids.txt
fetch_raw_xml.py
XML
raw_data.xml
acquire.py
JSON
games.json
JSON
games.json
Local
Local
Local
Local
KMeans
Clustering
app.py
ingest.py
JSON
games_clustered.json
model.py
JSON
games_featurized.json
featurize.py
Local
RDS
KMeans.pkl


S3
upload.py
download.py
BeefSack's GitHub
BoardGameGeek.com
BoardGameGeek.com
BGG XML API 2
boardgamegeek2
fetch_game_id.py
game_ids.txt
fetch_raw_xml.py
XML
raw_data.xml
acquire.py
JSON
games.json
JSON
games.json
Local
Local
Local
Local
KMeans
Clustering
app.py
ingest.py
JSON
games_clustered.json
model.py
JSON
games_featurized.json
featurize.py
Local
RDS
KMeans.pkl

Model Evaluation





Thank You!

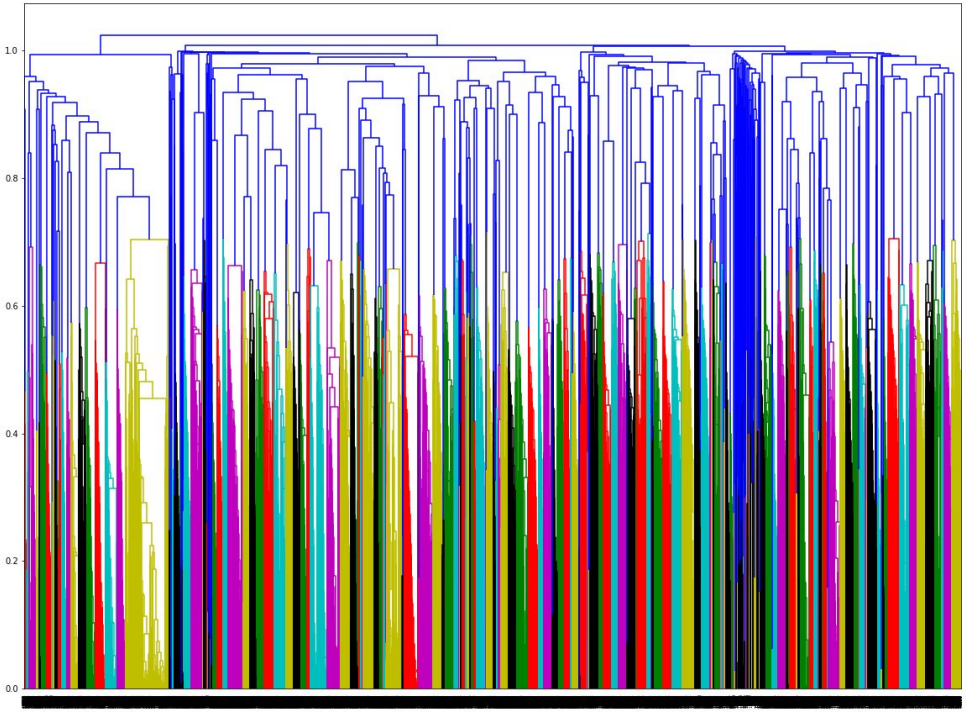
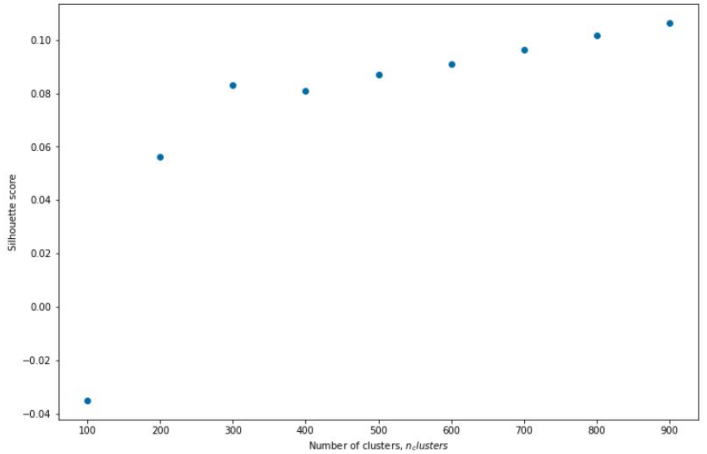
Kristiyan Dimitrov

2019 - 2020: Master of Science in Analytics, Northwestern University
2015 - 2019: Measurement & Attribution Product Specialist, Google
2011 - 2015: BA in Mathematics at the American University in Bulgaria

CV, GitHub

Appendix

Agglomerative Clustering



- Slower to Train
- Lower Silhouette Score