JIMMY GARDNER METIS 21’

SURFING FOR DATA

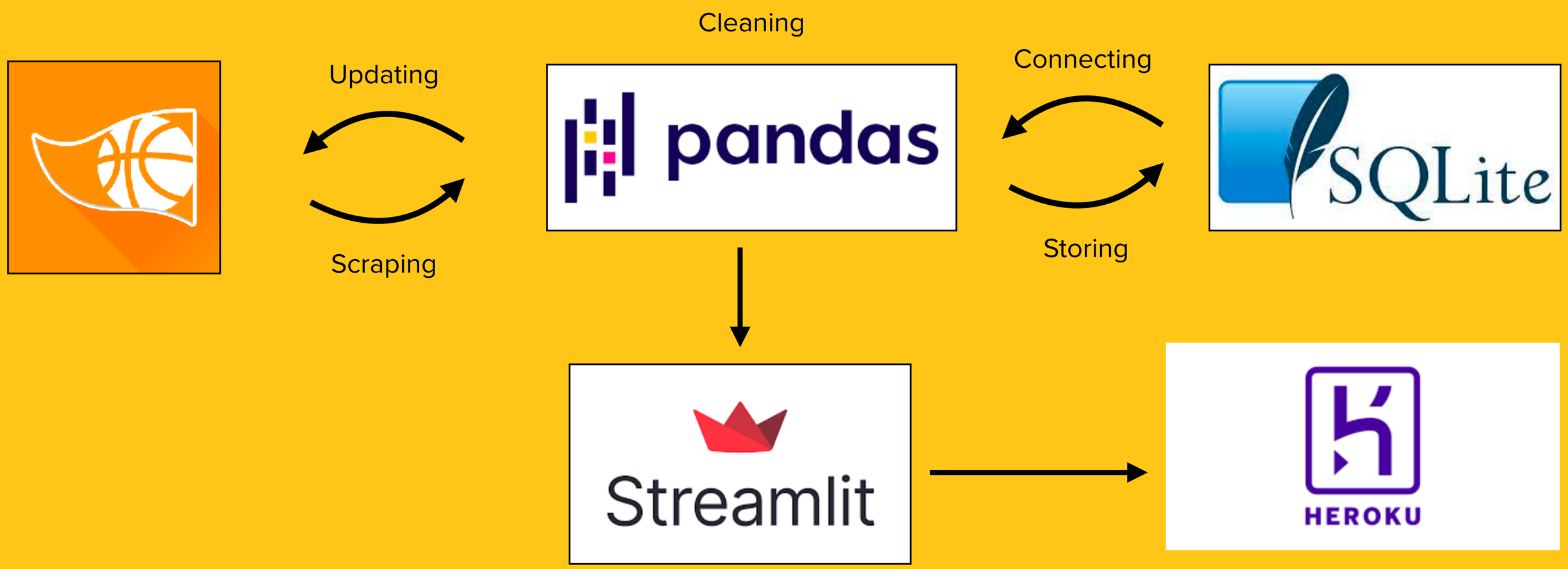
End to End Data Pipeline for WNBA and NBA Statistics

END TO END PIPELINE

GOALS

- Create a deployable data application
- Allow user interactivity with application interface
- Establish dynamic data flow that is easily updatable and integrate into application
- Build regression model for predicting Win Shares
- Modular Programming

PIPELINE

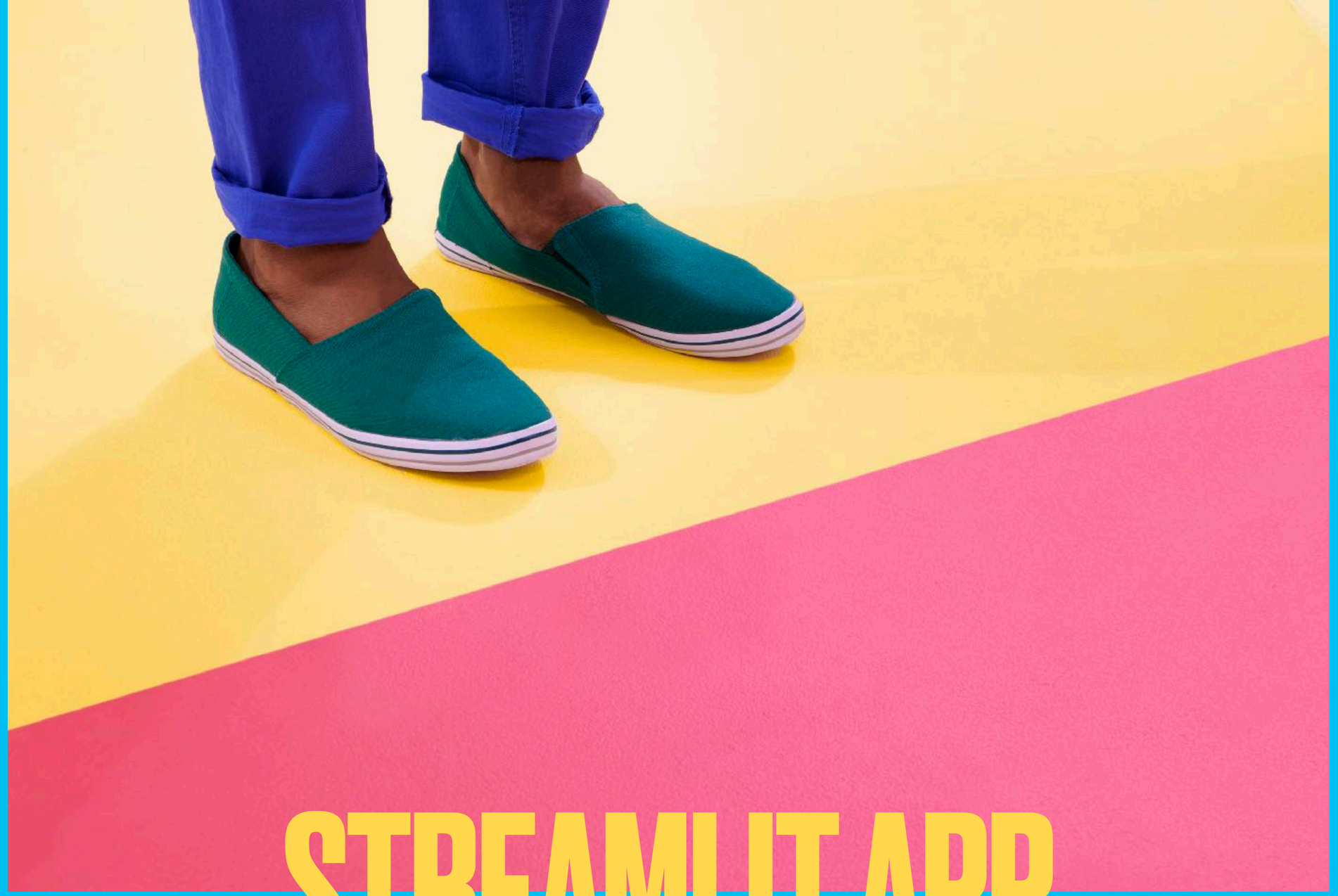

Cleaning
Updating
Scraping
pandas
Connecting
Storing
SQLite
Streamlit
HEROKU

PIPELINE STOPPAGES

HOW TO ENSURE PIPELINE IS OPERATING CORRECTLY

- Implement Unit Testing
- Be aware of any changes to html script from source site
- Update Scraping script to allow for retrieving only active players data
- Make sure database is populated

RESULTS



STREAMLIT APP

Lets look at the Streamlit application

FUTURE WORK

- Build Regression Model and implement in Streamlit application
- Transition from modular programming to object oriented programming
- Try building out application in Flask
- More data points
- More plots and interactivity