An aerial photograph of a desert landscape, likely in the southwestern United States, showing a prominent fault line running diagonally across the terrain. The terrain is arid, with sparse vegetation and distinct geological features. The title text is overlaid on the upper portion of the image.

Real-Time Data Pipeline & Dashboard for US Earthquake Analytics

By Nicholas Bronson

Agenda

1. Introduction & Problem
 2. Data & Pipeline Setup
 3. App Demonstration
 4. Conclusions & Next Steps
-

1. Introduction & Problem



Introduction: Earthquake Basics

- Tectonic plates build tension and eventually slip past each other over an area known as a fault line
- Earthquakes occur in waves, and vary drastically in strength, known as magnitude
- Earthquake magnitude is on a logarithmic scale, and an increase of 1 point can represent a dramatic difference in strength



Problem

Idea Catalyst: A CNN article asserted that the Blanco Fracture zone off the coast of Oregon has seen an uptick in earthquakes, some of which had a magnitude of 5.0+

Problem: Is the number of earthquakes in Oregon and off its coast increasing? Is there a way that we can easily investigate this?



2. Data & Pipeline Setup



Data

- 869141 rows of data with 22 features from US Geological Survey Quakes
 - Features include: magnitude, depth, magnitude error, latitude and longitude, depth, id, and depth error
- Data is pulled daily using an API, processed, and uploaded



Pipeline Setup



Data Pulled from
USGS using API



Data ingested
into DB Browser
SQL database



Data
preprocessing
and analysis
using pySpark



Further
preprocessing
using Pandas



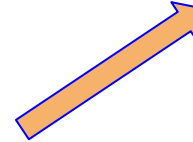
Data exported
uploaded to
GitHub and
SQL



Streamlit Share
page deployed,
using data from
GitHub

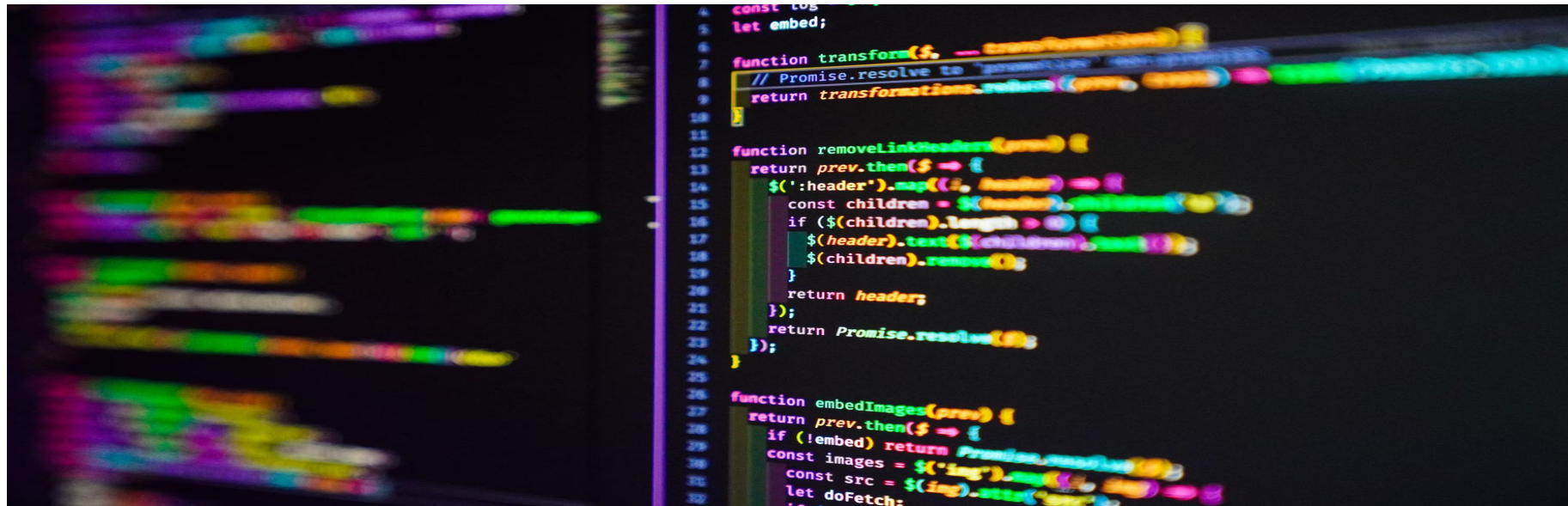


Real- time data
acquired using
API, processed
with one-click



3. App Demonstration

https://share.streamlit.io/bronsonnh/streamlit_repo/main/nick-app.py



4. Conclusions & Next Steps



Conclusions

- Oregon's increase in earthquakes seems modest, however a deeper look may be necessary
 - Perhaps newsflow on earthquakes was sensationalized
 - The vast majority of earthquakes in the database were under the detectable level, while it is interesting to notice these, it seems that studying trends around earthquakes of magnitude of 2.5+ may be more impactful
-

Next Steps

- **Shift storage from local machine to the cloud**
- **Increase level of automation**
- **Build additional features, improve aesthetics of web page**
- **Build a model that can predict number of earthquakes and strength of earthquakes in a given month**



Thank you!