Where does all the money go?

For project 2, Sonny and I(Jon) will be going through the financial assets in US$billions of the top banks in the world. We found this data here: <https://www.relbanks.com/worlds-top-banks/assets>

We will also be using waybackmachine to look at older asset numbers from previous balance sheets. Our exact data can be found below. We will web scrape the data using python and then add that data to PostgreSQL. Then we will proceed to the data analysis and visualization portions.

Data Analysis:

We will display a variety of charts using plotly. The first chart used will be a bar graph. This bar graph will show the number of banks in each country at each time increment. There will be a dropdown menu to go through each time increment. If we are high on time, we will make a line graph for each country, showing the rate of change in number of top banks through a time series line graph. There will be a dropdown menu to go through each time increment.

Query: Select Count(bank), country Where userinput=time increment groupby country

Query: Select Count(bank), time series where userinput=country groupby time series

Next, we will show another bar graph which shows the total assets in each country at each time increment. There will be a dropdown menu to go through each time increment. If we are high on time, we will make a line graph for each country, showing the rate of change in total assets through a time series line graph. There will be a dropdown menu to go through each time increment.

Query: Select Sum(Assets), Country where userinput=time increment groupby country

Query: Select Sum(Assets),time increment where userinput=country groupby time increment

Then we will have a list of the top 10 banks and their corresponding details for each time period. There will be a dropdown menu to go through each time period.

Query: Select Bank name, country, assets, time period where userinput=timeperiod orderby assets desc limit 10

Lastly, we will have time series line graphs which show the change in assets in individual banks. There will be a dropdown menu to cycle through different banks.

Query: Select assets, time period where userinput = bank

We will be using Animate on Scroll(<https://michalsnik.github.io/aos/>) as our new JS library to have all graphs on one page.

Our pending question to explore is to see if there is a trend in money moving between different countries. Should a company try to move their assets elsewhere? Are some countries’ banking sectors in trouble?

Data

<https://web.archive.org/web/20120626023412/https://www.relbanks.com/worlds-top-banks/assets> which covers March 2012 balance sheets

<https://web.archive.org/web/20121225122536/https://www.relbanks.com/worlds-top-banks/assets>

which covers September 2012 balance sheets

<https://web.archive.org/web/20131111012357/https://www.relbanks.com/worlds-top-banks/assets>

which covers March 2013 balance sheets

<https://web.archive.org/web/20140209050456/https://www.relbanks.com/worlds-top-banks/assets>

which covers September 2013

<https://web.archive.org/web/20140701123856/https://www.relbanks.com/worlds-top-banks/assets>

which covers March 2014

<https://web.archive.org/web/20150512003339/https://www.relbanks.com/worlds-top-banks/assets>

which covers December 2014 as September was not available

<https://web.archive.org/web/20151104133716/https://www.relbanks.com/worlds-top-banks/assets>

which covers June 2015

<https://web.archive.org/web/20160504071841/https://www.relbanks.com/worlds-top-banks/assets>

which covers December 2015

<https://web.archive.org/web/20161120235852/https://www.relbanks.com/worlds-top-banks/assets>

which covers June 2016

<https://web.archive.org/web/20170515081439/https://www.relbanks.com/worlds-top-banks/assets>

which covers December 2016

<https://web.archive.org/web/20180205184632/https://www.relbanks.com/worlds-top-banks/assets>

which covers June 2017

<https://web.archive.org/web/20181224170514/https://www.relbanks.com/worlds-top-banks/assets/>

which covers December 2017

Steps:

1. Web scrape Data
2. Clean Data
3. Export Data to SQL
4. Made flask app and connected to SQLite database
5. Made app routes for our 5graphs and 1 table
6. Made html page and linked it to our flask app and our js file
7. Made divs for our 5 graphs
8. Made all 5 graphs in javascript
9. Make table in html
10. Add rows from javascript to html table space using arrays made for table
11. Make app routes for dropdown lists
12. Make dropdown lists for all 6 graphs/table
13. Add change setting functions and implement them in js/html
14. Clean data for more errors if found while exploring visualizations
15. Add x and y axis titles to all graphs

Still to do:

Expound more on details for graphs/what is each graph showing

Aesthetics work

Add animate on scroll fade functionality

Explore our visualizations and find a storyline/data driven decision

Research external evidence to backup our claim. Add this evidence to html page in an aesthetic way.

Proposal: IRS Criminal Investigation Applied Analytics Division employees approaching the Strategy Director to start a IMF taskforce to combat tax evasion and tax fraud. The data gathered would be the start of an initiative to gather data about the assets of banks, both public and private, and use this data while juxtaposed with gdp data. This normalized data should show any large deviations going forward if bank assets largely exceed expected values given gdp and current market health. These deviations would prompt investigations from governments involved. Overall, this finished data would promote many of the IMF’s values while also stopping corporate greed on a larger scale and a healthier market for all.

Steps to do: Make graphs of China and USA GDP growth over same time periods with plotly and data right into java script. Compare these graphs to USA and China assets over time

Presentation Points:

Exploratory bubble- look into trends in banking assets in USA

Our data was comprised of only publicly banks and not all big banks were tracked or tracked consistently throughout data gathering processed. Found evidence of correlation between banking assets of country and GDP. May be able to use this data to look for discrepancies in tax filings. GDP and assets should have a relationship. If accounting for trends in investment and savings, a deviation in assets in regard to GDP suggests that money is being moved in ways it should not. Too little money would suggest money being moved overseas to take advantage of lower taxes or lower fees. Too much would suggest money coming in from overseas. It is in the IRS’s best interest to keep money where it should be. Scaling up our research will allow for better and more complete data. A key fault in this theory is that the current ratio would be one already corrupted by existing evasion and fraud. However, establishing a link will stop evasion on a higher scale.

We also propose a joint initiative with the IMF. As money must go somewhere or come from somewhere, a global effort will allow for discover money flow trends and will further help with identifying current culprits. A global initiative will allow the IRS and other likeminded organizations in other countries to track down and rectify pending cases to a higher degree. In the long run the extra resources devoted towards this project will be pennies compared the money siphoned illegally from the markets.

Slides:

Intro slide with task assigned

Downfalls of initial data slide

US assets vs US GDP slide/ discussion of slide numbers

Showcase of why its so important to take all bank assets: graph of Goldman vs BoA

Discussion of possible benefits and costs from expanding nationwide data.

Discussion of how this venture would be more fruitful on a global scale

Examination of China/other example country which shows opposite trend of US data