

## CS-6630 - Data Visualization

### Brazil: Credit operations in public sector and the Human Development Index

<https://github.com/fredericosar/credit-operations>

Frederico Rocha - Anil Kumar Konasale Krishna - Anil Kumar Ravindra Mallapur  
[fred@cs.utah.edu](mailto:fred@cs.utah.edu) - u0810018 - [anil.konasale@utah.edu](mailto:anil.konasale@utah.edu) - u0939372 - [anilmr@cs.utah.edu](mailto:anilmr@cs.utah.edu) - u0942509

#### 1. Background and Motivation

Brazil has entered in a fiscal crisis, not as intense as the one Greek is going through, but equally worrying. In his latest research, Leno Rocha, an analyst for the Brazilian National Treasury, showed how the number of loan pleas made by federative entities evolves during a fiscal year. He was able to prove that the diffusion phenomenon is determinant on the dynamics of the pleas. Furthermore, his research demonstrated how the number of credit operations fits in the Gompertz curve. The theory shows excellent results when compared to the practical data from 2002 to 2014. His work demonstrates that large financial volume of credit is concentrated in a small number of operations, enhancing the predictability and the performance capacity of debt managers in the Brazilian National Treasury.

We believe this work is especially interesting because it allows us to make predictions on Brazilian financing pleas. It also allows researchers to understand the important consequences of the federal debt, such as the moment of the credit impact in the social and Brazilian domestic macroeconomic scenario. Moreover, it is possible to categorize the finality of the credit operations (e.g. Health, Infrastructure, *etc*) and use the Human Development Index to analyse if the HDI determines who is more likely to take the loans.

Thus, the reason for us to implement a visualization for this research and data, is to analyse the credit operation environment that surrounds Brazil's development and show how it might be possible to connect it to an humanized development index, and provide a better understanding of how the credit operations affect quality of people's life. We can also use visualization to show the predictions of these pleas using the Gompertz curve.

Finally, the visualization can be used in the management of human resources. Understanding the demand of pleas creates the possibility to foresee the amount of work that will be necessary and efficiently allocate the workers that analyze the pleas for subnational loans. This efficient allocation improves the quality of spending on personnel expenses in the administration of Brazil's public finances.

#### 2. Project Objectives

This project has primarily 7 objectives. They are described below:

- a. Demonstrate that the number of credit operations made by brazilian federative entities fits the Gompertz curve.

Question to be answered: Is the simulated data accurate?

Learn and accomplish: Learn more about the Gompertz curve and prove Leno's theory right.

- b. Categorize the type of creditors to check where the loaned money is coming from.  
Question: From where the loaned money is coming from?  
Learn and accomplish: Who invest in Brazil as a creditor? How frequent they invest?
- c. Demonstrate that the most part of the loaned money is concentrated on a small number of credit operations.  
Question: How the loan money is being distributed?
- d. Categorize the finality of the credit loans between infrastructure, health, education, and safety.  
Question: Why cities and states are asking for loans?  
Learn and accomplish: Learn why Brazilian federative entities ask for loans.
- e. Compare the HDI evolution during the years of 2002 to 2014 and analyse if the number of credit operations (loans) has any role in the index changes.
- f. Try to describe facts about the requested loans.  
Learn and accomplish: Is there any strong reason the federative entities are asking for money?
- g. Show how the HDI changed during the 12 years that are being analyzed.  
Question: Are the credit operations related to HDI changes?  
Learn and accomplish: How the credit operations affects the Brazilian HDI.

### 3. Data

The credit data comes from the Brazilian Treasury Secretariat<sup>1</sup>. It provides the information about the credit operations in the Brazilian public sector from 2002 to 2014. The data is offered in a CSV (comma separated value) file.

The data contains the Human Development Index (HDI) for Brazilian cities and states. It comes from the United Nations<sup>2</sup> and from Google Public<sup>3</sup>. The data in both sites are available in a table format.

The Brazilian latitudes, longitudes and areas were located in multiple places. We decided to use the ones offered by Globo<sup>4</sup> and the official Brazilian Data website<sup>5</sup>.

1. <http://sadipem.tesouro.gov.br>
2. <http://www.pnud.org.br/atlas/ranking/Ranking-IDHM-Municipios-2000.aspx>
3. <http://www.google.com/publicdata>
4. <http://app.globoesporte.globo.com/futebol/mapa-das-torcidas-no-facebook/data/fbz.json>
5. <http://dados.gov.br/>

#### Data Example

The credit data contains 14 fields. However, only 7 fields are useful for this work. They are presented with sample data below.

City	State	Credit Finality	Type of Creditor	Creditor Name	Amount in US Dollars	Date
------	-------	-----------------	------------------	---------------	----------------------	------

Goiânia	Goiás	Infrastructure	Public National	Caixa	\$80,922.21	09-10-2013
Rio de Janeiro	Rio de Janeiro	Health	Public National	Banco do Brasil	\$3,140,016.00	06-07-2011

The HDI data contains 6 fields. Again, we are only interested in the fields shown below.

City	Human Development Index
Goiânia	0,715
Rio de Janeiro	0,716

#### 4. Data Processing

The credit operations data is provided in Portuguese and will be translated to English. As it already has well-defined fields, it will not be necessary to do a lot of cleanup, just the exclusion of some fields. However, the data file needs to be parsed to JSON to simplify the manipulation using D3. For that, we will use free online converting tools such as CSV to JSON web-site<sup>6</sup>. Using this credit data, we generate the Gompertz curve, we categorize the type of creditor in three categories (Public National, Private National, International) and the credit finality in four categories (Infrastructure, Health, Education, and Safety). Also, as the amount of the credit loan is based on the currency of the country in which the creditor is located, we will convert the amount to U.S. dollars.

The data for HDI is very straightforward and won't need any cleanup nor derivation. This is also valid for the geographical data.

Lastly, the data processing will be done on-the-fly using Javascript on the loading of the page. As all the data will be well-defined and offered as a JSON this will enable fast update on the page.

6. <http://www.csvjson.com/csv2json>

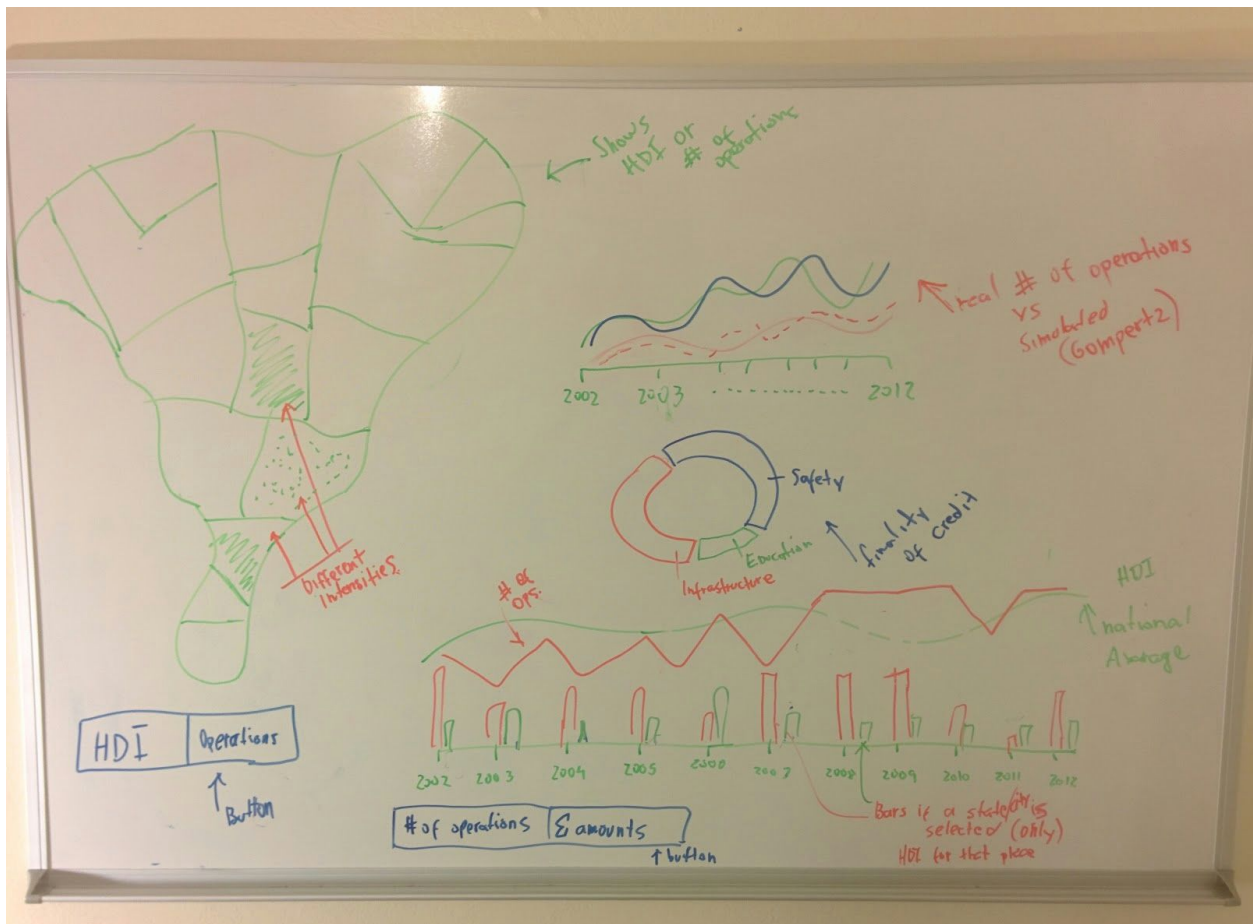
#### 5. Visualization Design

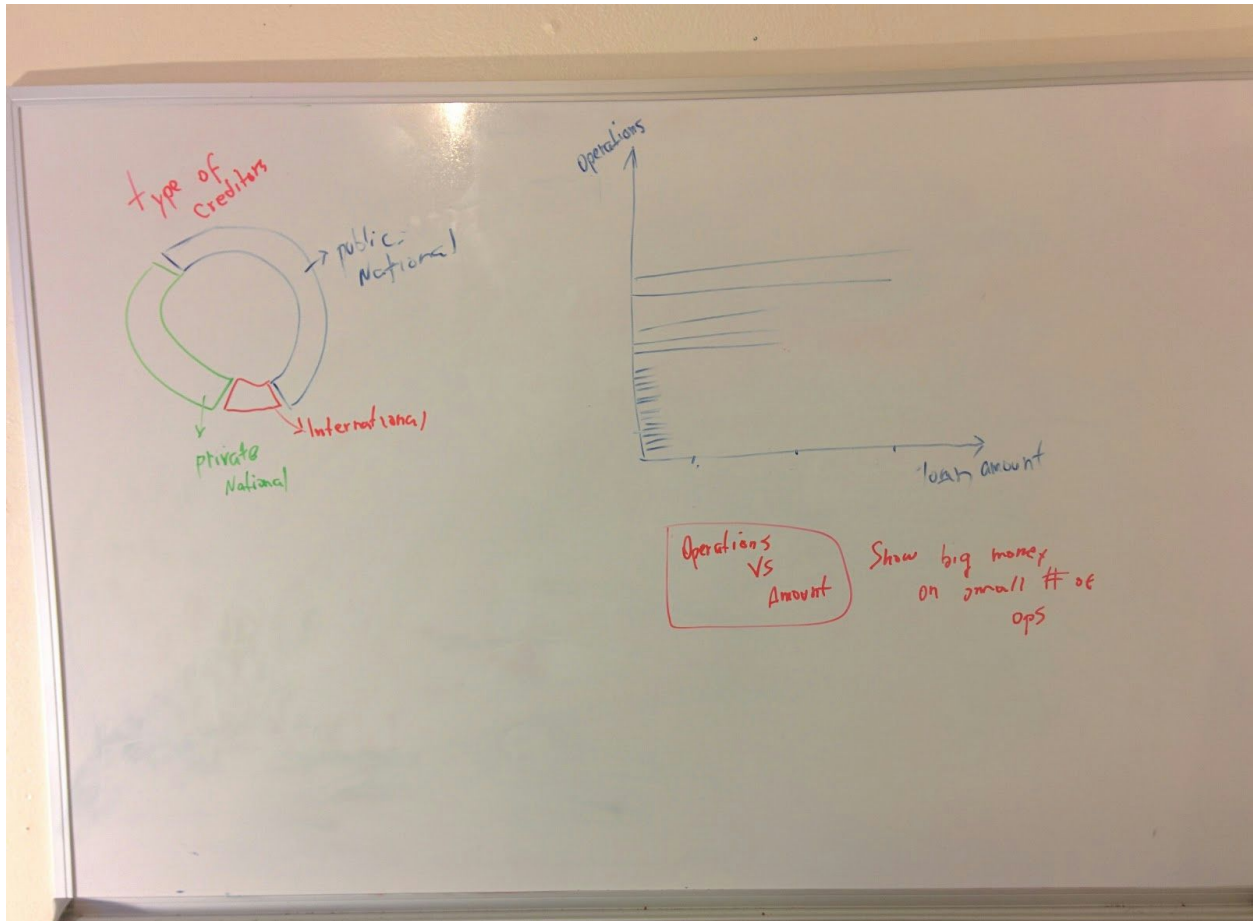
##### a. Prototype One

This design has 6 main components that are described above.

- Brazilian Map (map) - The map will show the average HDI or an information about the credit operations (total value vs number of operations) for each Brazilian state. It was chosen because a map can provide a very good connection between locations and the information that is trying to be shown.
- Total credit operation (line charts) - The number of credit operations will show the theoretical number and the practical data for the total amount credit operations. This chart is very important because it will give the user the possibility of analysing future simulated data. Also, it is a great way to show that the Gompertz curve works excellently for this case.

- Finality of credit (donut chart) - The finality chart is where we will show the reason each federative entity used to explain why they needed the loan. Also, it can also be used in the attempt to make a connection with the HDI. This chart was chosen because it is a nice way to show parts of elements that is will be summed to a total.
- Type of creditor (donut chart) - This chart will show from where the loan money is coming from. It is nice to check if the loans are coming from a third country or if it is coming from a national public or private creditor. A donut chart was chosen for the same reason the finality of credit uses it.
- HDI vs Total number of operations line/bar chart - This is one of the main charts when thinking about showing an information for a general public. It shows on the same graph the HDI averages and the number of credit operations. It can also show data from an specific state if any of the 26 is selected.
- Operations versus Operation amount - One of Leno's claim is that there are a very big amount of money in a very small number of operations. This graph will be useful to corroborate with this data and see how much the states or cities are normally asking for.





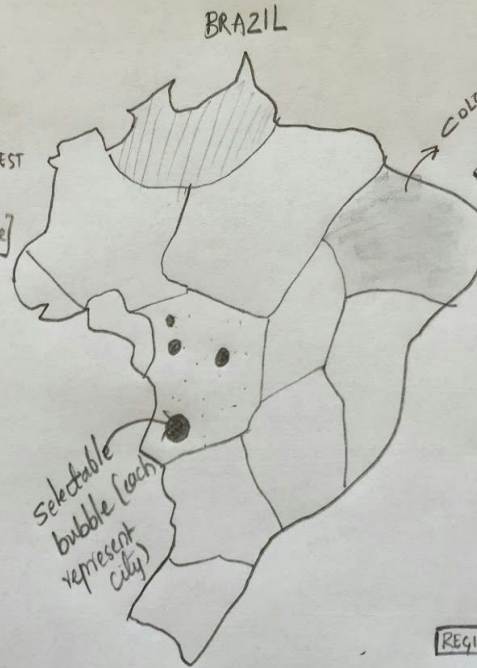
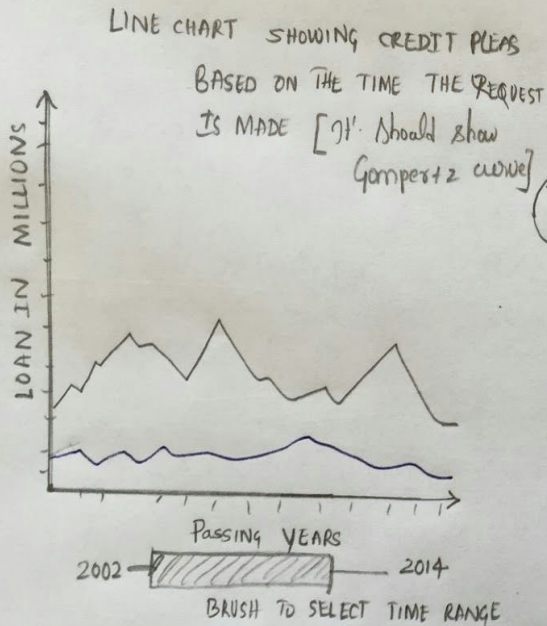
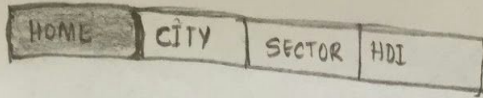
## b. Prototype Two

In this design, we would like to show how the credit data that we have been using can be visualized to show the overall credit requests. The above design shows the main visualization page with Brazilian Map with regional boundaries. As visualization is started, the credit requests made by all cities in a particular region would be considered. The color intensity of each boundaries shows at that particular point of time who made the most amount of credit request with respect to the total amount of money requested. A user can select the city map instead of regions, which would represent cities as bubbles and size indicates the amount of money requested by that city. A city for example which has hosted FIFA's soccer world cup 2014 would have made credit request for larger amount of money to compensate the infrastructure cost. The line chart towards the left side shows the progressive line over the span of 2002-2014 considering all the requests by all cities. If multiple cities are selected from map, one can visualize and compare two line charts in the same chart. One being the total request made and other being credit requests made by those cities. The fact sheet beside shows concrete evidence of why a particular time had seen more amount of credit requests for infrastructure or other verticals. For example election time, world cup time. The second page of Visualization

lets user search for a city, and shows the line chart divided upon different verticals like infrastructure, education, job market for which credit was requested from year 2002 till 2014. There is a stacked bar graph which shows the total amount of credit requested for a particular year and the amount of credit sanctioned as not all credits gets approved. If a user clicks on particular bar, the simple bar graph beside shows the HDI ( Human Development Index for that city ) for that year and updates the two pie chart below which shows the distribution of approved loans for different sectors and share of creditors who sanctioned that loan.

# BRAZIL: CREDIT OPERATIONS AND HDI.

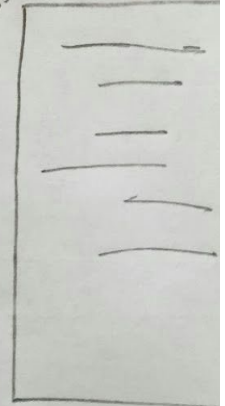
TABS



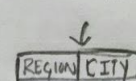
COLOR INTENSITY  
SHOWS THE  
REGIONS  
CREDIT  
PLEAS.

What developments  
happened to  
GDP &  
HDI  
to R  
an  
ax

FACTS



TOGGLE BUTTON

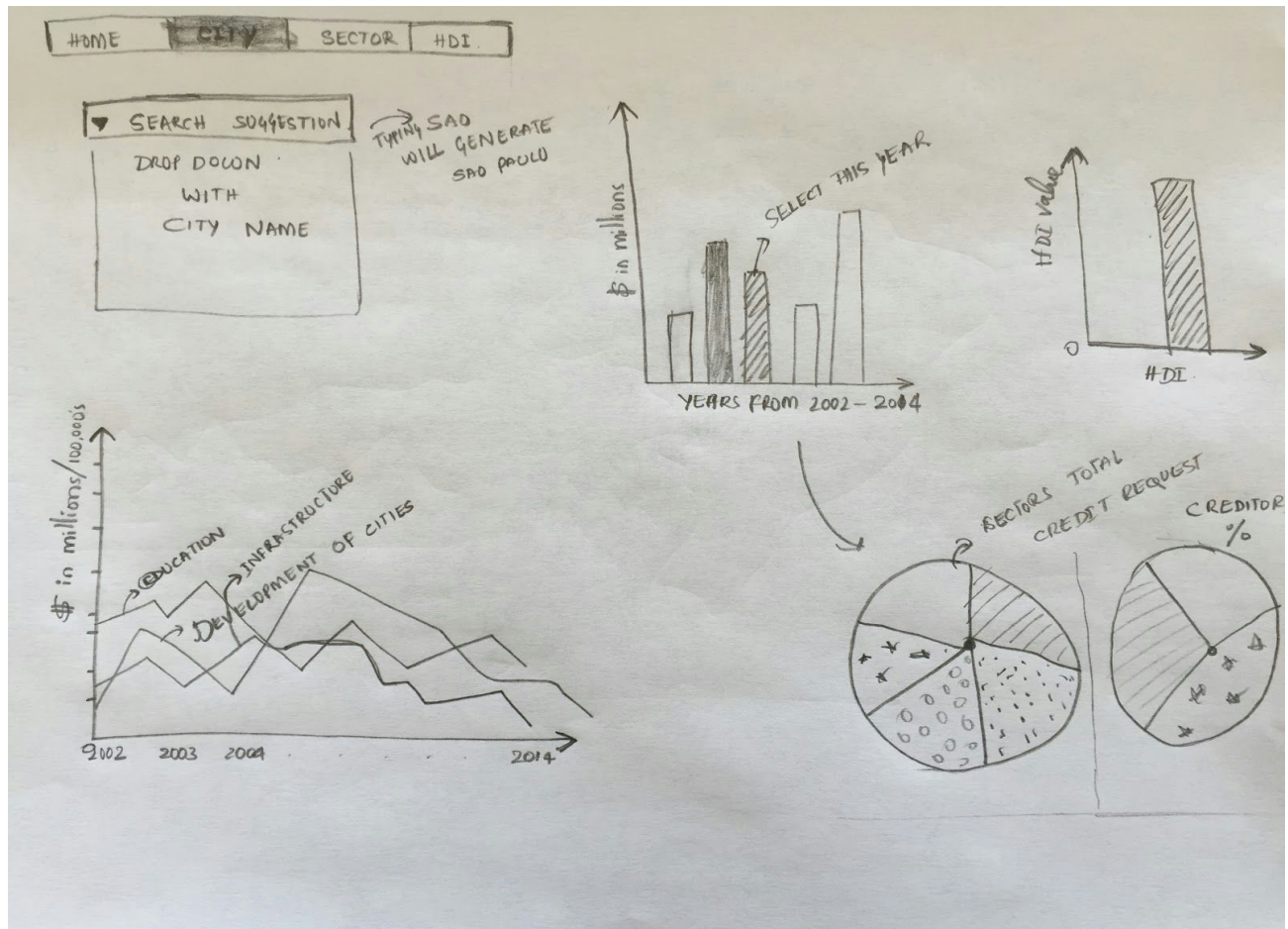


[If] City, then  
each city with  
bubble showing  
the size & color  
Proportion to  
Credit plea



Note: ① When city is clicked then the line chart will have blue line which shows that city's credit requests.





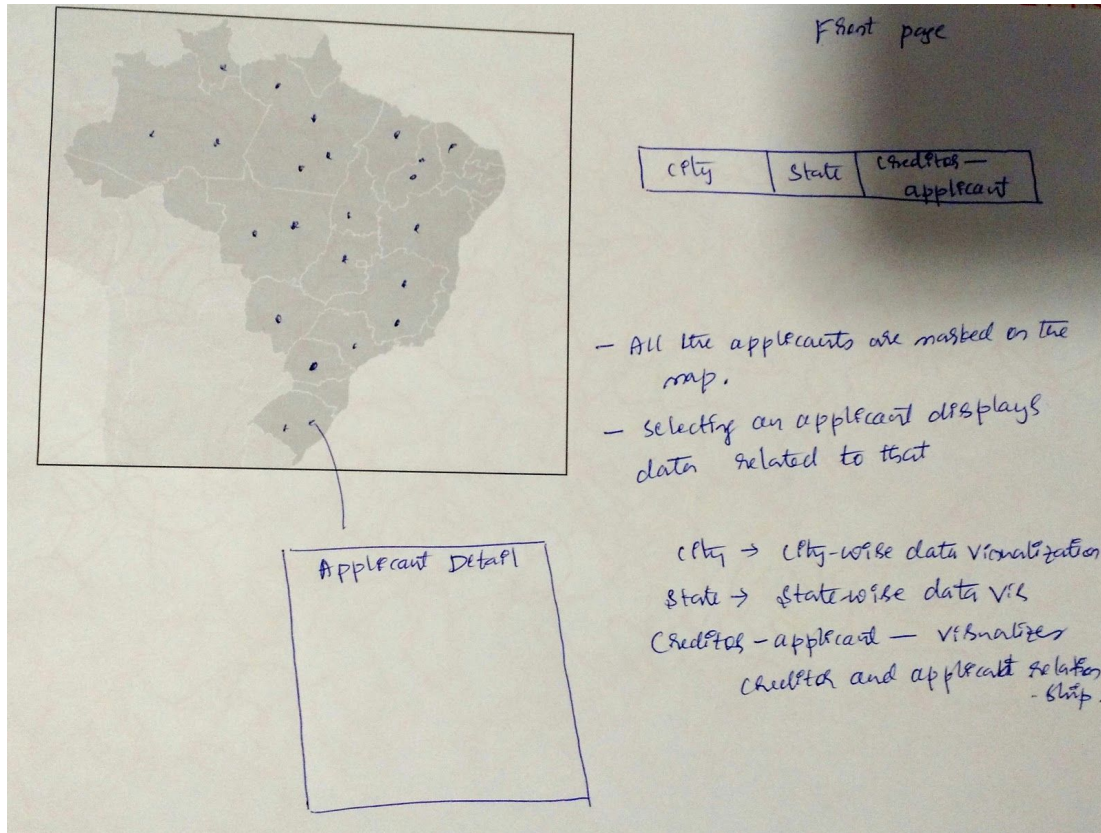
### Prototype Three

The design visualizes the credit data with each visualization emphasising a particular information. Below prototype consists of a front page and three visualization.

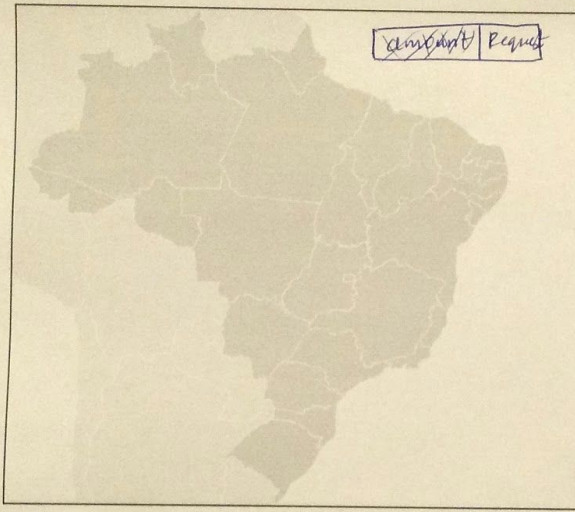
- Front Page - It marks all the loan applications on the Brazil Map with state boundaries. Selecting an applicant displays complete details pertaining to that particular applicant in details. This Visualization is very helpful in getting individual applicant details. Front Page also consists of buttons to city, State and Creditor-Applicant visualizations.
- City Visualization - Brazil Map marked by boundaries is colored with different colors and intensities to indicate number of applicants or amount of loan requested in each cities based on the selection (Amount/Requests). Selecting a year, visualizes data for that particular year. Bar graph on the right side indicates the city-wise data. Line graph visualizes credit finality over cities. One can find determine a fast growing city by looking at the line graph. City with highest amount of loan requested in infrastructure projects, industrial investment etc.



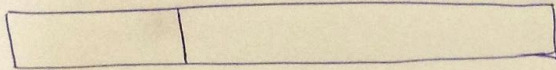
- **Creditor-Applicant Visualization** - This establishes relationship between applicants and creditors. Creditor Bar Graph displays amount of loan requested from each creditors. This can be used to find the creditor with highest and lowest amounts of requested loan. Applicant Bar Chart displays amount of loan requested by each applicant with respect to a particular creditor. A creditor without higher loan amount might indicate that he is more reliable than one with lower amount.



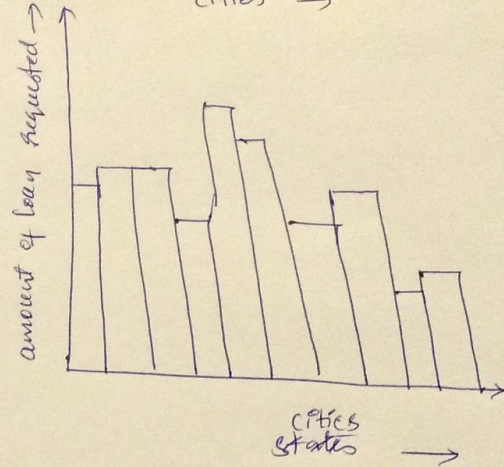
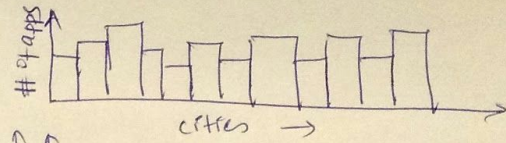
# State-wise Visualization



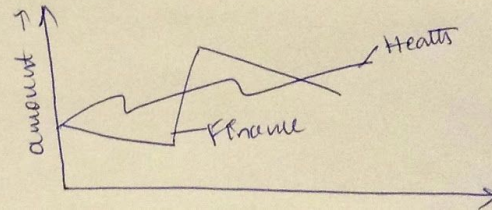
Brazil Map - showing amount & loan requested in each cities

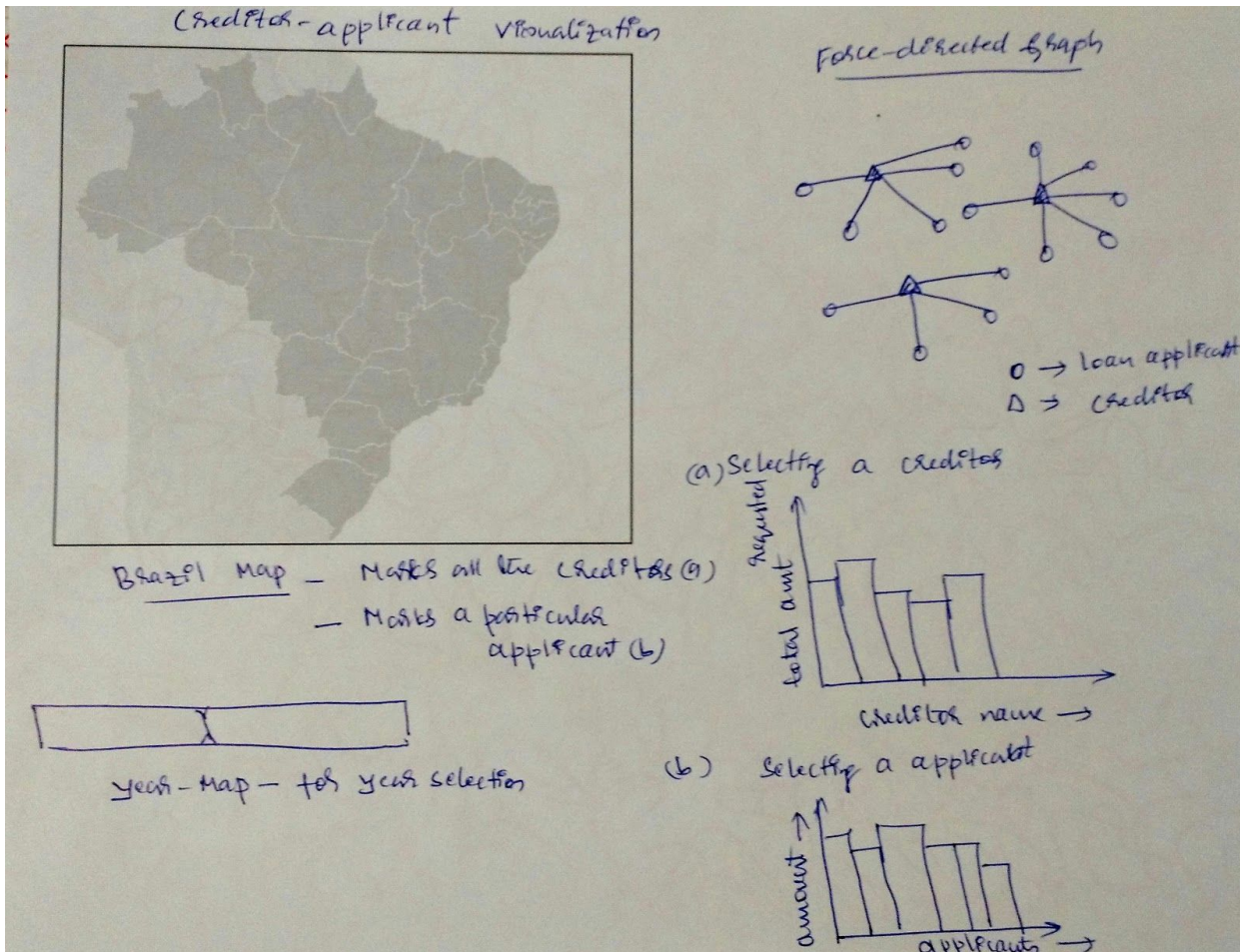


Year Bar - year selection



Total amount of loan requested in each states.

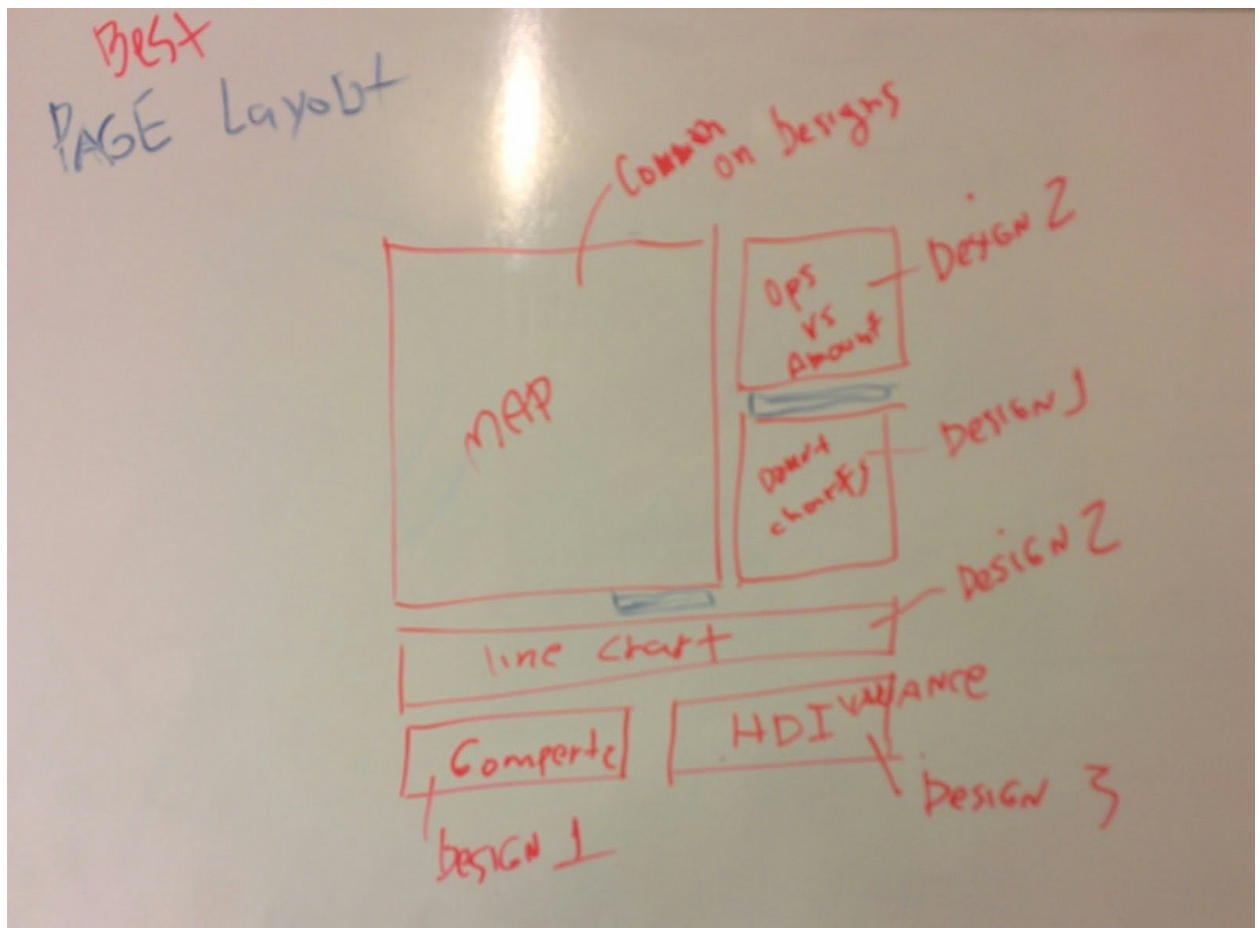




### Best prototype (features from previous prototypes)

We think the our design will follow the diagram shown below:





The best features (the one will be implemented):

Brazilian map - Shown on both designs.

Operation vs Amount requested - As shown on design 1 (and 2).

Donut charts for credit type and creditor type - As shown on design 1.

Line charts for HDI and number of credit operations - As shown on design 2 (and 1)

Gompertz curve - As shown on design 1.

HDI variance - As shown in design 1 (and 3).

## 6. Must-Have Features

- Brazilian interactive map that allows zooming on states. It will be colored using a scaled based on the number of operations for a state or city or the total amount of the credit loaned. It can also be based on the HDI of the state.
- A line chart showing the real historical number of credit operations and the number of credit operations simulated using the Gompertz curve.

- c. A line chart showing the HDI average and the total number of operations. On the same chart, if a state is selected, a bar chart will be added showing the average HDI and the number of operations for that state.
- d. A donut chart showing the amount of operations for each credit finality category.

## 7. Optional Features

- a. A donut chart showing the amount of operations for each type of creditor.
- b. A bar chart showing the evolution of the HDI for the country or for each state.
- c. Visualize the graphs using a timeline.

## 8. Project Schedule

What?	Who?	Week #44	Week #45	Week #46	Week #47	Week #48	Week #49
Translate Data	Frederico						
Categorize Data	Frederico						
Convert Data to JSON	Anil M						
HTML structure	Anil KK						
Structure MVC for Javascript	Anil KK and Anil M						
Map Visualization	Anil M						
Timeline Visualization	Anil M and Frederico						
City Search Feature	Anil KK						
Stacked bar graph	Anil M						
Line Charts	Frederico						
Donut Charts	Anil KK and Anil M						
CSS, Prettify, Fact Sheets	Everyone						