

# Patent Patterns

Explore the World of Patents

## The Patent Industry

- Patents are a means of rewarding innovation and creativity
- In 2010, IP intensive industries accounted for 34.8% (\$5.06 Trillion) of U.S. GDP
- Process of filing a patent is highly complex with lots of rules and regulations
- Not much research has been done around patent applications versus granted patents

## Motivation

- Uncover data driven insights from the publicly available patents data
- Better understand which patents get through USPTO
- Investigate gender bias in patent applications and grants
- Deliver clean data and insights to guide future research and improvements in the field of patents

## Process

### Data Collection

- Scraped data from USPTO using PHP and Unix
- Processed and stored data using AWS EC2/S3

### Data Cleaning

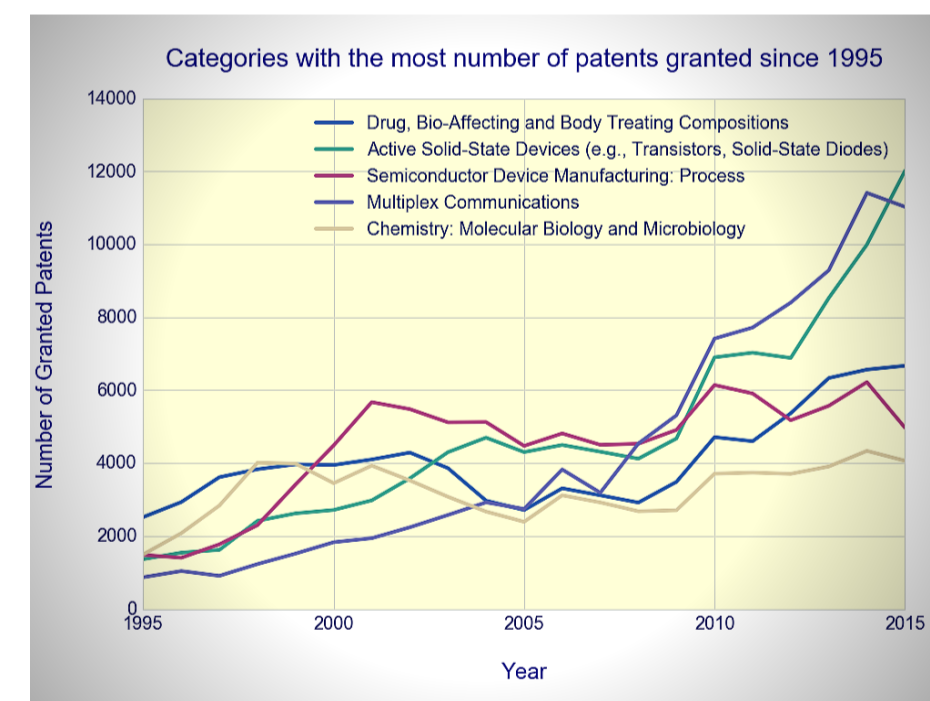
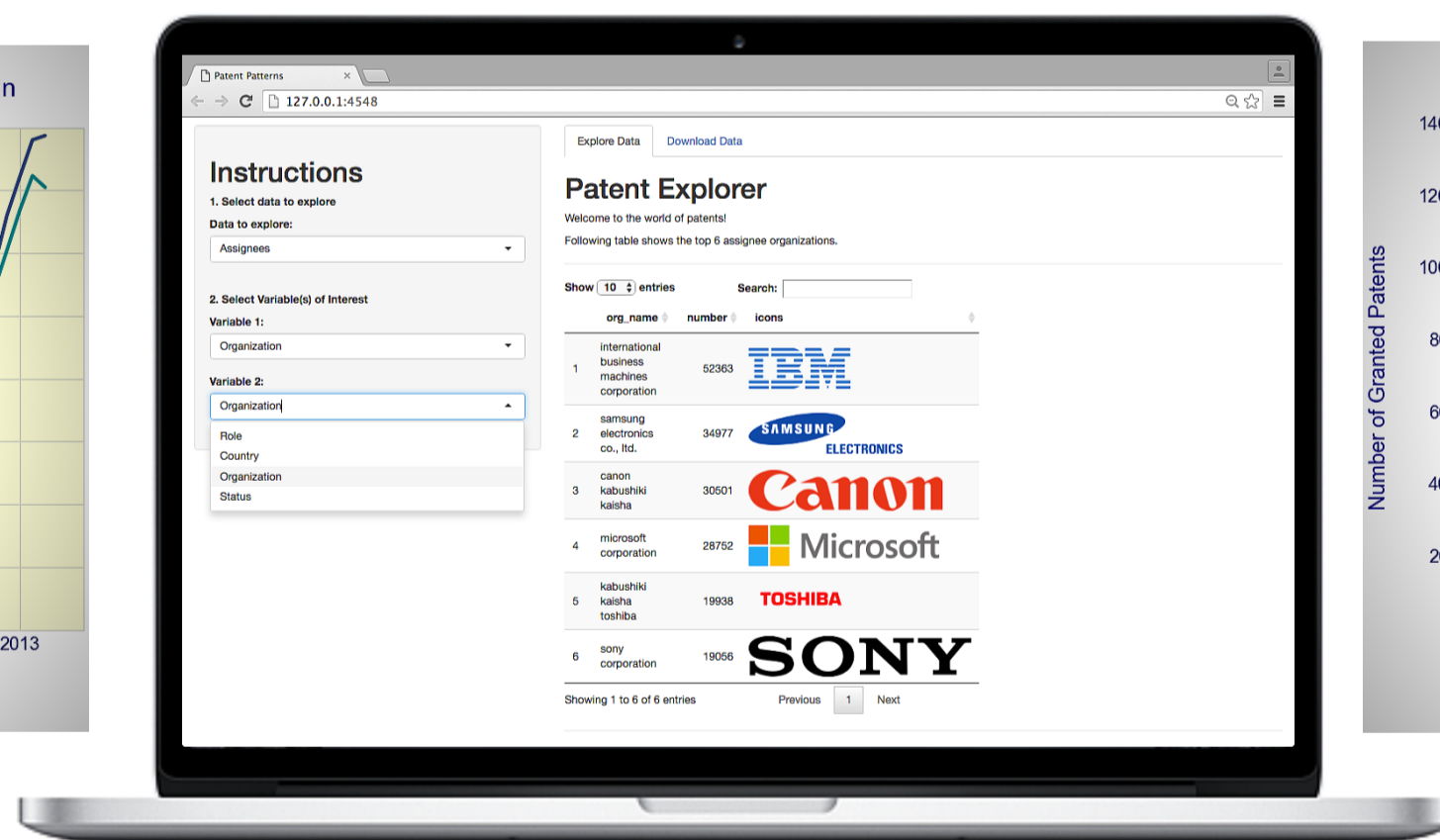
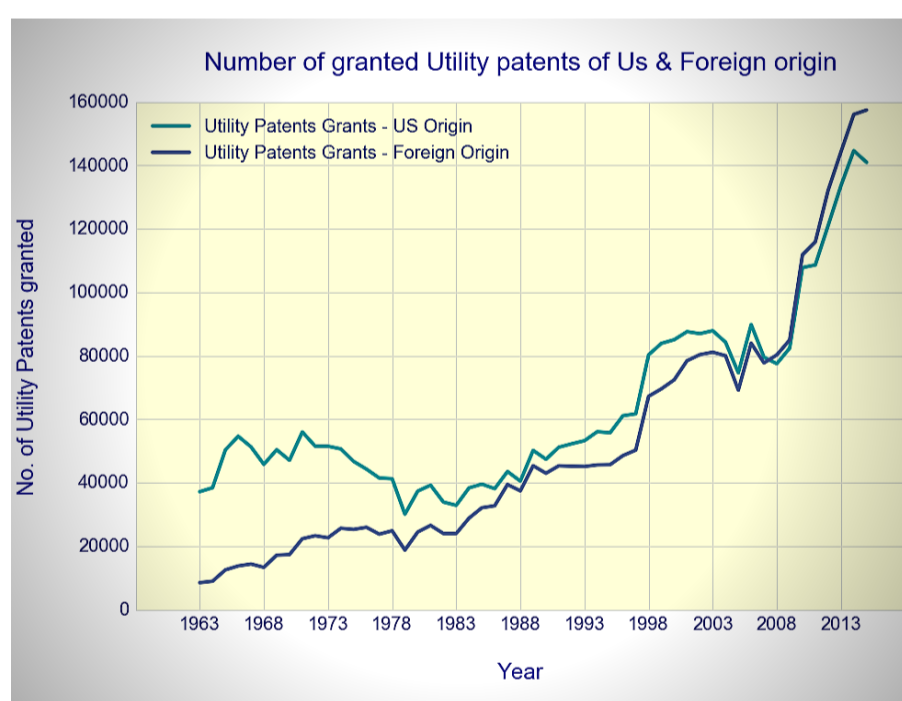
- Merged Datasets
- Cleaned data sets using Python

### Data Analysis

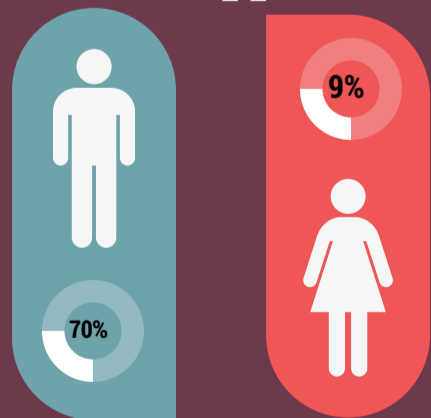
- Performed Exploratory Data Analysis using Python / R

### Results

- Static Visualizations using Matplotlib
- Built a visualization app using R (Shiny)



### Patent Applications



### Granted Applications



## Issues

- Unavailability of whole dataset
- Inconsistencies in the data
- Complex patent filing rules in the USPTO
- Missing citations data

## Future Work

- More in-depth feature engineering
- Tune classifier to predict issuance or rejection
- Make our app more interactive

### Application Assignees

From U.S.A

41%

24%

From Japan