World GDP Prediction & Analysis

Aakash Moghariya, Ruturaj Patel and Vikash Yadav

Introduction

What is gross domestic product (GDP)?
- Primary indicator used to gauge health of a country's economy
- expressed in terms of percentage of improvement or degradation

Factors affecting GDP of a country
- Working force, production, oil prices and many other indicators can impact GDP of a country in either positive or negative way.

Our Project

- Our project uses feature selection algorithm such as Pearson correlation to determine the factors that affect the GDP of a country the most.

- Using up to 10 important factors, we train different regression model to predict the GDP of a country for future years.

Data Cleaning & Integration

- Data in CSV and XML were cleaned using tools such as Spark, Python and Pandas.

- Combined 33 different indicators affecting major countries of the world

- Resolved country names in different files and separating them from the data of continents and other regions

- Ensured mono temporal data along with proper scales

- Levistine distance was used to resolve the UN country name with official country name and ISO alpha 3 country code.

Datasets :

UNdata — A world of information
WORLD BANK GROUP
OECD — BETTER POLICIES FOR BETTER LIVES
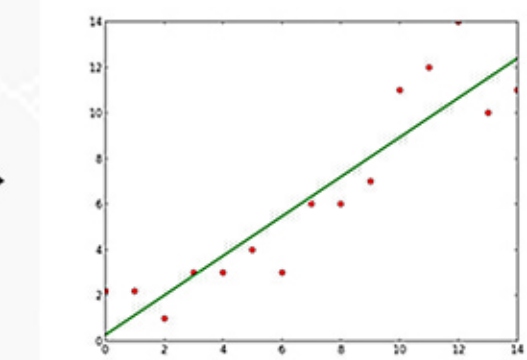KPMG

Machine Learning Pipeline

Clean Data

Pearson Correlation

$$r = \frac{\sum XY - \dfrac{\sum X \sum Y}{N}}{\sqrt{(\sum X^2 - \dfrac{(\sum X)^2}{N})(\sum Y^2 - \dfrac{(\sum Y)^2}{N})}}$$
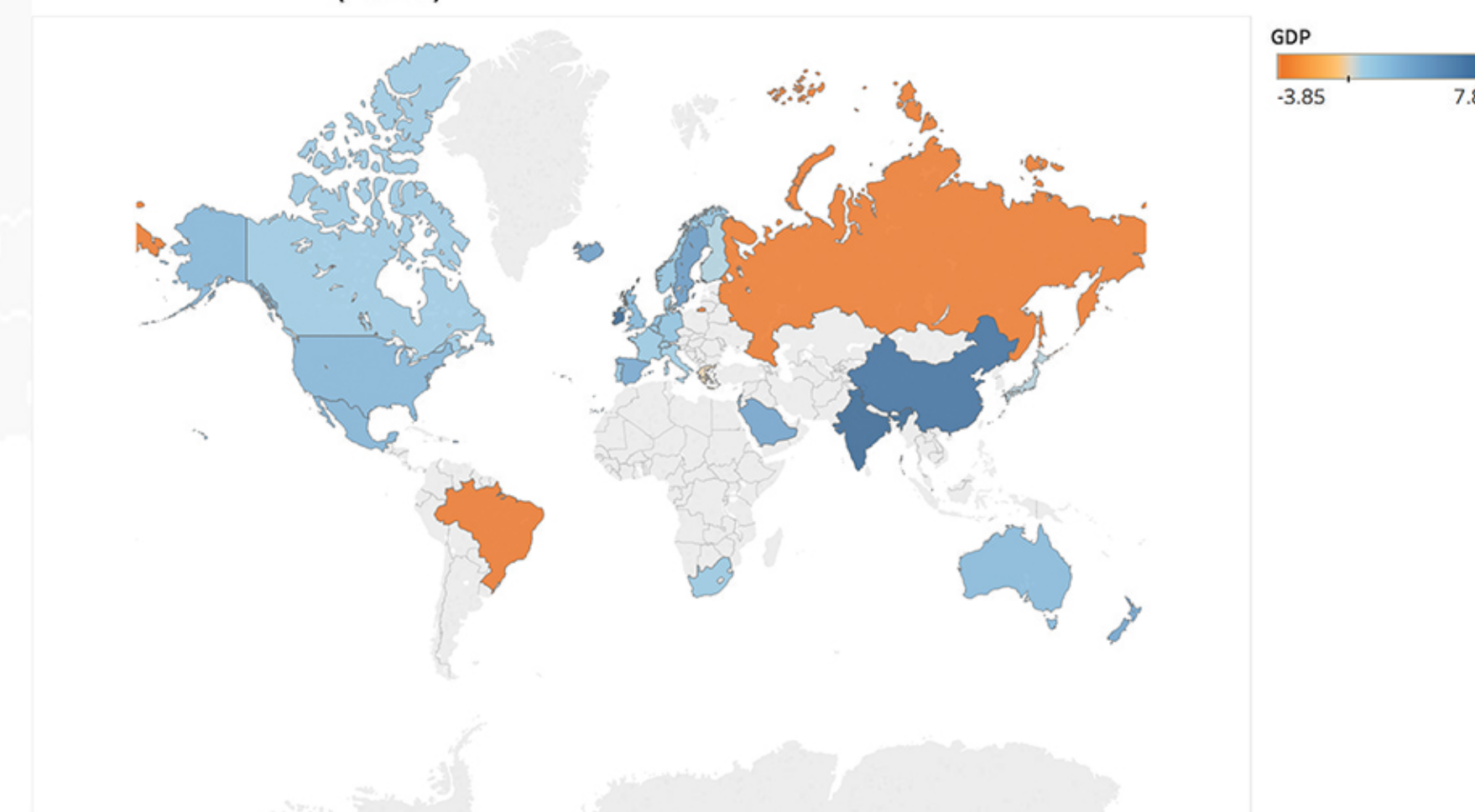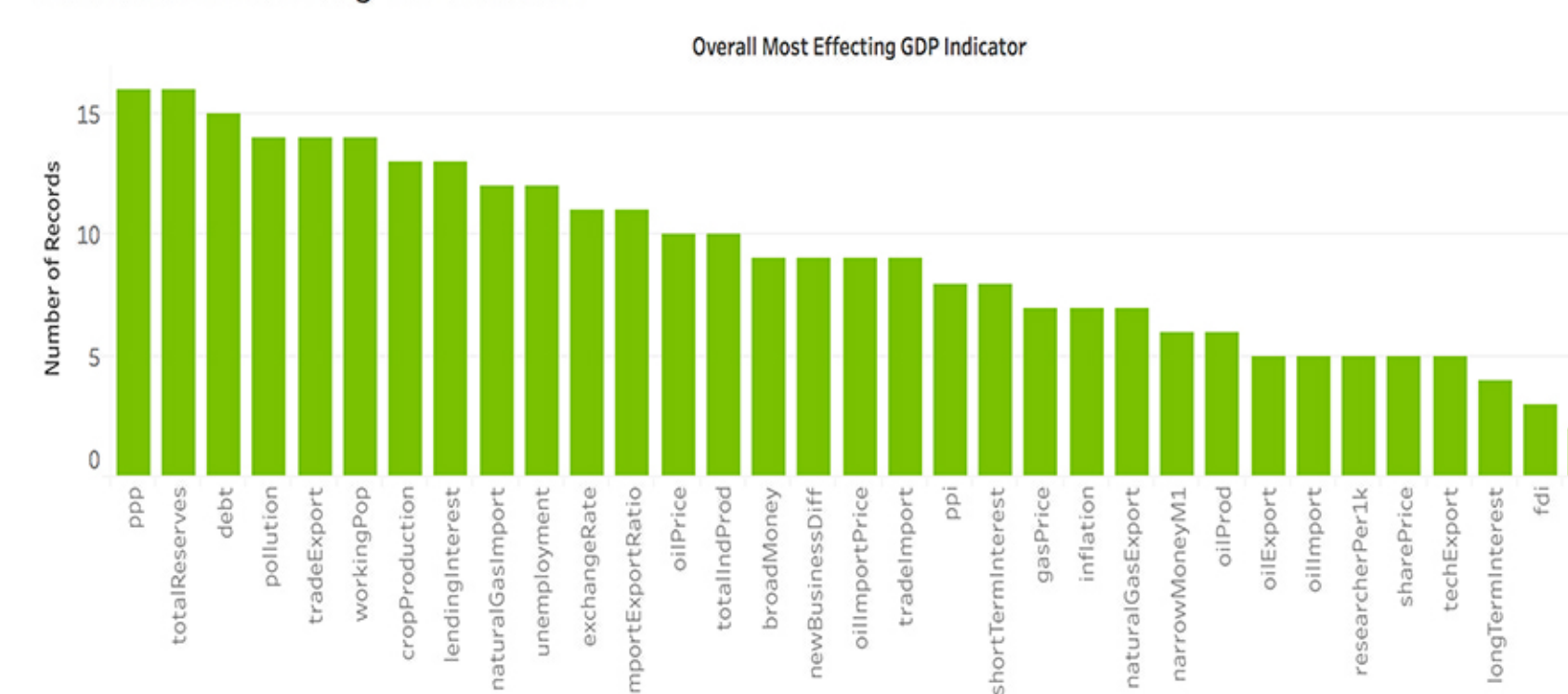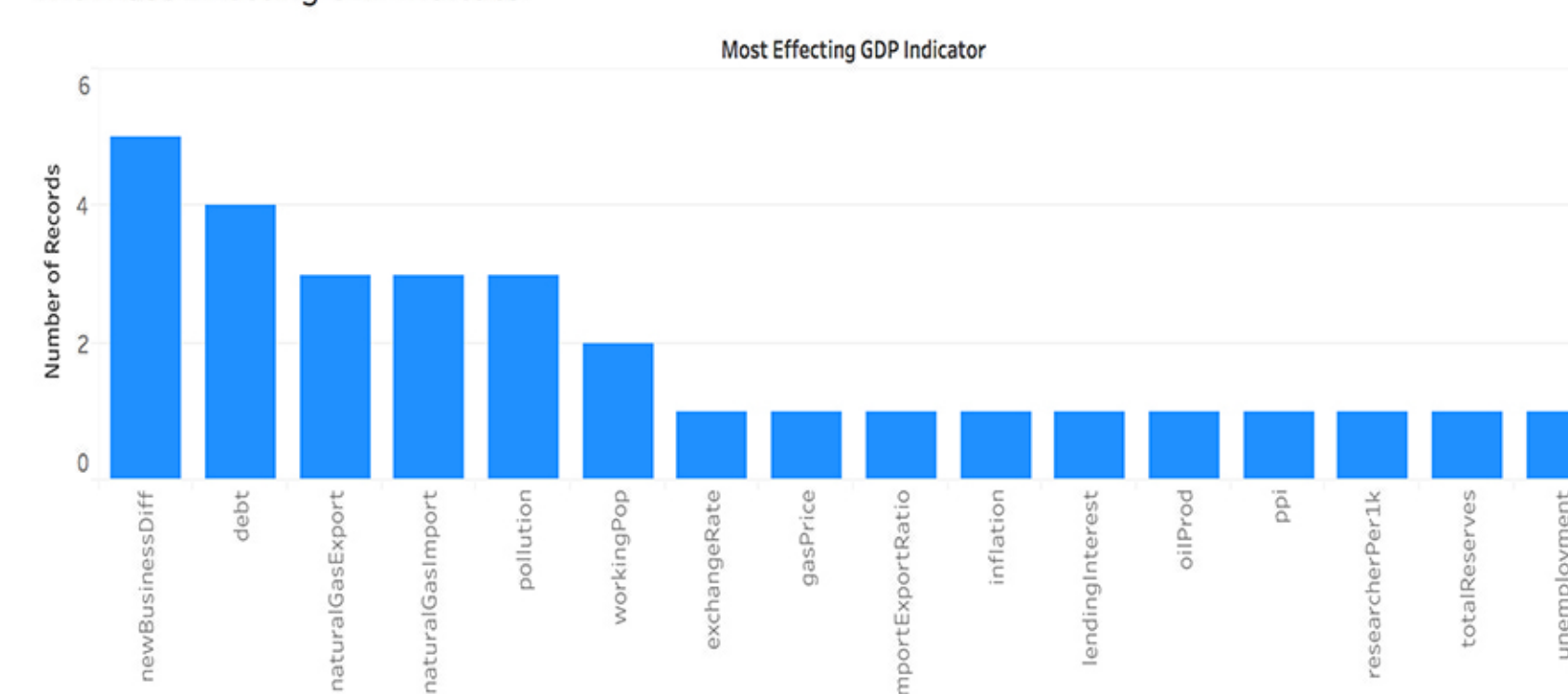
Regression

Technology Stack:

APACHE SPARK
python
pandas
$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$
tableau

Results

GDP Over the Year (2015)

GDP -3.85 — 7.81

2015 GDP(Actual V/S Predicted)

Measure Names
Actual
Predicted
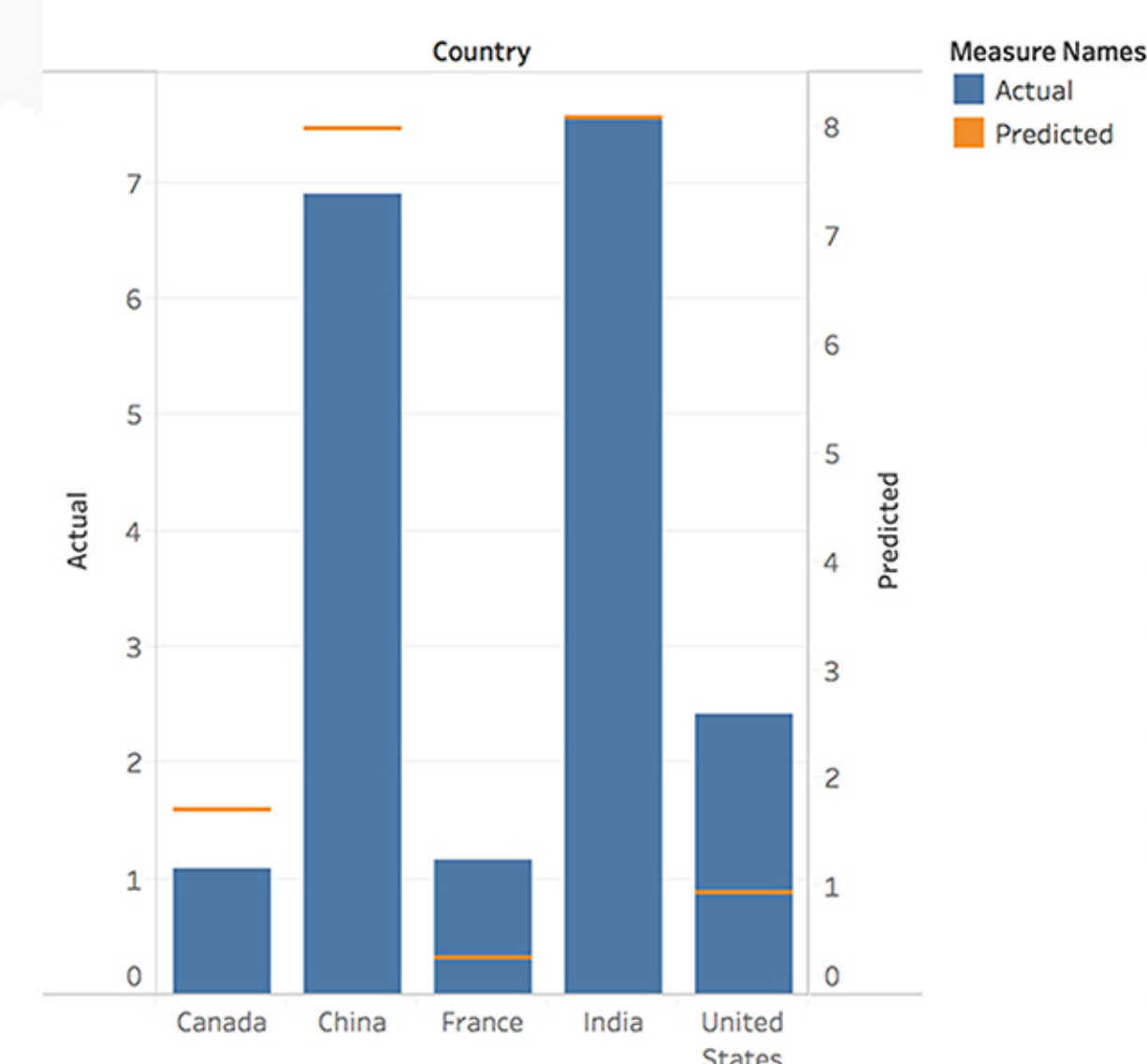
Country	Year	Actual	Predicted
Canada	2014	2.4728926	1.7333236
Canada	2015	1.0782688	1.7216044
India	2014	7.2434717	8.080998
India	2015	7.5701304	8.1063842
United States	2014	2.4277956	1.9496427
United States	2015	2.4259705	0.9735468
China	2014	7.2685132	8.9075309
China	2015	6.9	8.0029189
France	2014	0.2594941	0.3283883
France	2015	1.155967	0.3660394

Overall Most Effecting GDP Indicator

The Most Effecting GDP Indicator

Correlation of Indicator vs GDP growth rate