****

**Predictive modeling on business data (Financial ratios) for loan processing**

|  |  |  |
| --- | --- | --- |
| Batch/Roll No | Name of the Student | Specialization |
| PGDM6/1910 | Akash Gupta | Finance |
| PGDM6/1903 | Aayush Bansal | Finance |
| PGDM6/1950 | Sonal Roy | Finance |
| PGDM6/1918 | Ayushi Garg | Mkt |
| PGDM6/1932 | Oshin Srivastava | HR |
| PGDM6/1953 | Sunanda Sarkar | Finance |

**Dataset link (GITHUB):** https://github.com/AkashGupta1997-colab/business-data/raw/main/BUSINESS\_DATA.xlsx

**REPORT LINK(GITHUB):**

**Google colab Code:** https://colab.research.google.com/drive/1lAmcVMUtIZBxrnZ9NISM3KcpxS5nsL3R?usp=sharing

**Date of submission: 5th Jan 2021**

**Name of Module Descriptor: Prof. Savio Saldhana**

**PROJECT OBJECTIVE:**

To Perform Predictive Modelling aimed at solving business problem.

**BUISINESS PROBLEM:**

1. Business problem that is identified for redressal using predictive Modelling –

To identify financial condition of a company that shows whether a bank should give loan to the company or not and analyze it using financial ratios.

1. Domain of the Business –

Banking sector/Financial firm, Business Loan

1. Reason for the Selection of the problem –

To enhance and improve the processes and procedures involved in a loan department of a bank.

1. Current Impact of the problem –

By using this model, banks will be able to identify whether a company would be able to repay the loan money or not. Also, this could help in avoiding future NPAs’.

1. Potential Positive Effects of Remedying the Problem –

* Analysis of financial ratios will lead to easier Identification of financial health of a company
* Process optimization loan sanctioning department
* Prevention of banks from the companies that can turn into a Non-Performing Asset (NPA).

**DATA IDENTIFICATION AND GATHERING**

We have identified 6 Data variables.

1. 5 Independent Variables i.e., D/E ratio, Operating Profit margin, ROCE, Net Profit Ratio, ROE
2. 1 dependent variable i.e., company’s financial condition.

**Source of data: NSE.com, Screener.com**

1. **Debt-to-Equity (D/E) Ratio**

The debt-to-equity (D/E) ratio is calculated by dividing a company’s total liabilities by its shareholder equity. These numbers are available on the balance sheet of a company’s financial statements.

The ratio is used to evaluate [a company's financial leverage](https://www.investopedia.com/ask/answers/040915/what-considered-good-net-debttoequity-ratio.asp).

​Debt/Equity= Total Liabilities / Total Shareholders’ Equity

A high debt/equity ratio is often associated with high risk; it means that a company has been aggressive in financing its growth with debt.

* The debt-to-equity (D/E) ratio compares a company’s total liabilities to its shareholder equity and can be used to evaluate how much leverage a company is using.
* Investors will often modify the D/E ratio to focus on long-term debt only because the risk of long-term liabilities are different than for short-term debt and payables.

1. **OPM (Operating profit margin)**

The Operating Profit Margin (OPM) ratio measures operational efficiency of a company and its pricing strategy. The OPM ratio indicates how much profit a company makes after paying for variable costs of production such as wages, raw materials, etc. It is also expressed as a percentage of sales and then shows the efficiency of a company controlling the costs and expenses associated with business operations.

Operating Profit Margin = Operating Income / Sales Revenue

A higher OPM ratio reflects efficiency of the business in procuring raw materials and converting them into finished products.

* A company's operating profit margin ratio tells you how well the company's operations contribute to its profitability. For instance, a company with a substantial profit margin ratio makes more money on each dollar of sales than a company with a narrow profit margin.

1. **Return on Capital Employed (ROCE)**

Return on capital employed (ROCE) is a [financial ratio](https://www.investopedia.com/ask/answers/062215/what-are-financial-risk-ratios-and-how-are-they-used-measure-risk.asp) that can be used in assessing a company's profitability and capital efficiency. In other words, the ratio can help to understand how well a company is generating profits from its capital.

ROCE= EBIT/ Capital Employed

​where: EBIT=Earnings before interest and tax

Capital Employed=Total assets − Current liabilities​

* ROCE is a metric for analyzing profitability, and potentially comparing profitability levels across companies in terms of capital. ROCE can be especially useful when comparing the performance of companies in capital-intensive sectors.

1. **Net Profit Ratio**

Net Profit Margin ratio, it establishes a relationship between net profit earned and net revenue generated from operations (net sales). Net profit ratio is a profitability ratio which is expressed as a percentage hence it is multiplied by 100.

Net profit ratio **helps to find out net profit earned in comparison to revenue earned from operations.** NP ratio helps to determine the overall efficiency of the business’ operations, furthermore, it is an indicator of how well a company’s trading activities are performing.

Net profit ratio = Net profit/ sales(revenue)

* A high ratio may indicate low direct and indirect costs which will result in a higher net profit of the organization.
* A low ratio may indicate unnecessarily high direct and indirect costs which will result in a lower net profit of the organization, thus reducing the numerator to lower than the desired number.

# **Return on Equity (ROE)**

Return on equity (ROE) is a measure of financial performance calculated by dividing [net income](https://www.investopedia.com/terms/n/netincome.asp) by [shareholders' equity](https://www.investopedia.com/terms/s/shareholdersequity.asp). Because shareholders' equity is equal to a company’s assets minus its debt, ROE is considered the [return on net assets](https://www.investopedia.com/ask/answers/070914/what-are-main-differences-between-return-equity-roe-and-return-assets-roa.asp). ROE is considered a measure of the profitability of a corporation in relation to stockholders’ equity.

ROE = **Net Income/ Average Stakeholders’ equity** employed in that period

* A high return on assets shows than the business was able to successfully utilize the resources provided by its equity investors and the company’s accumulated profits in generating income.
* Nonetheless, just like any other financial ratio, the ROE is more useful if it is compared to a benchmark such as the average ROE in the industry where the company operates or the company's ROE in the past years.

**MODEL BUILDING**

Cleaning a Data Set: We have created our own data, so that data cleaning is not required.

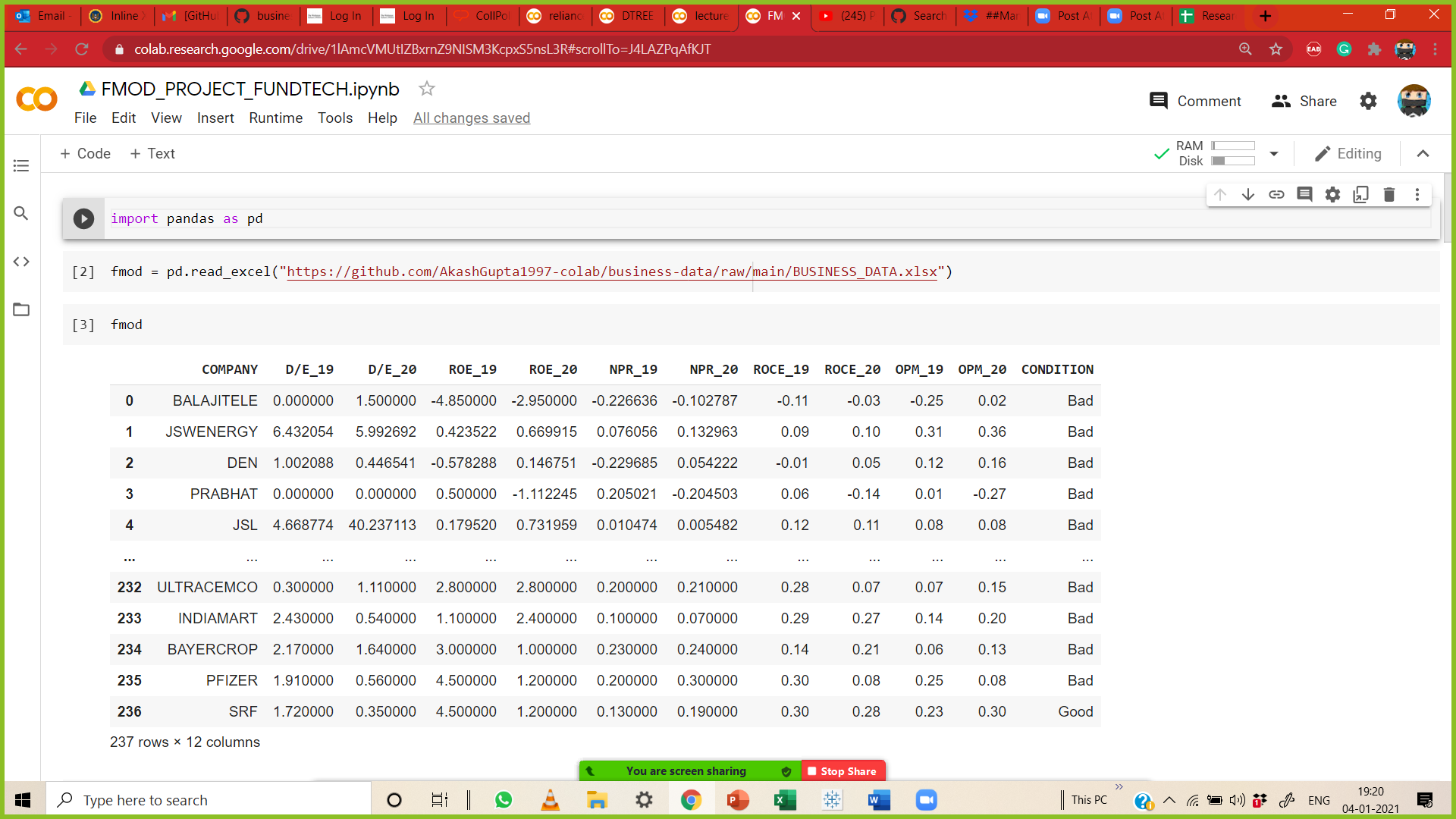
Dependent and Independent variables- After performing predictive modeling we found that few variables have very high p-value. And for that we have removed those variables. Our dependent variable is Condition column which state the financial condition of the company. And our independent variables are D/E\_ 20, ROE\_19, ROE\_20, NPR\_20, ROCE\_19, ROCE\_20.

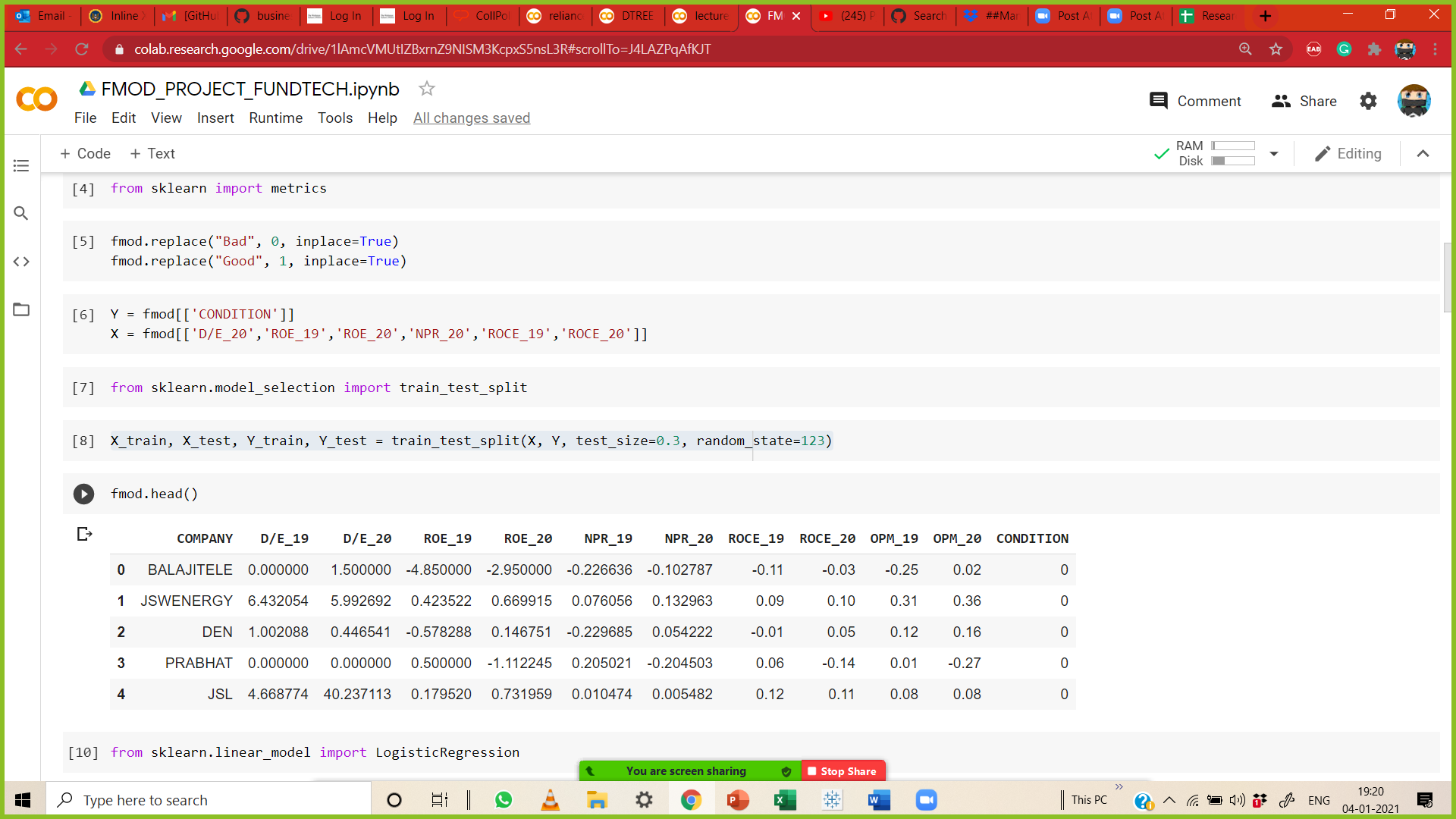
Data Partition- We have divide the whole data set into two part i.e. training and testing. 30% of data is consider for test the model and remaining 70% data is used for train the model.

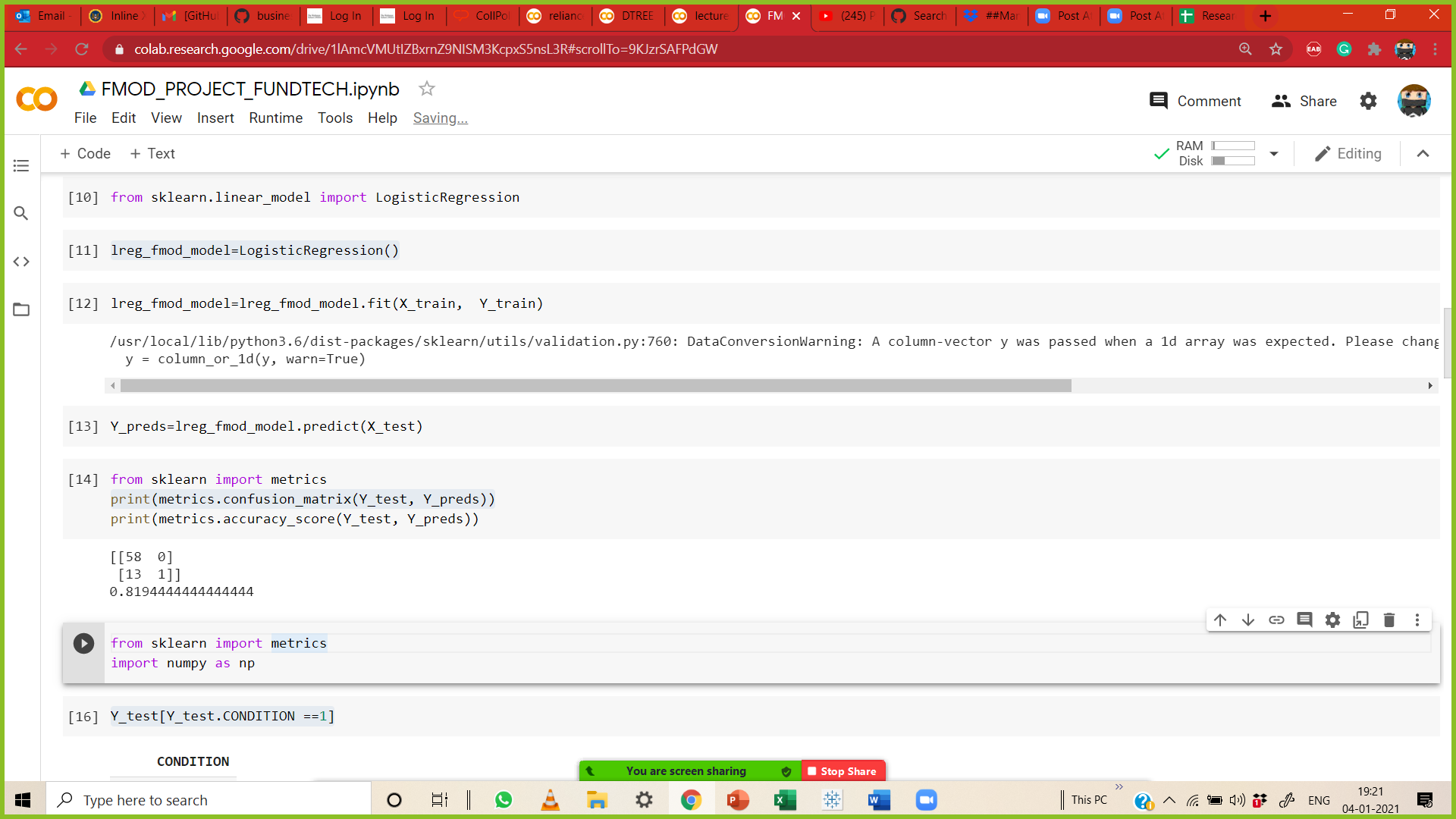
Model selection and reason for selection: Following models we have used for predictive modeling

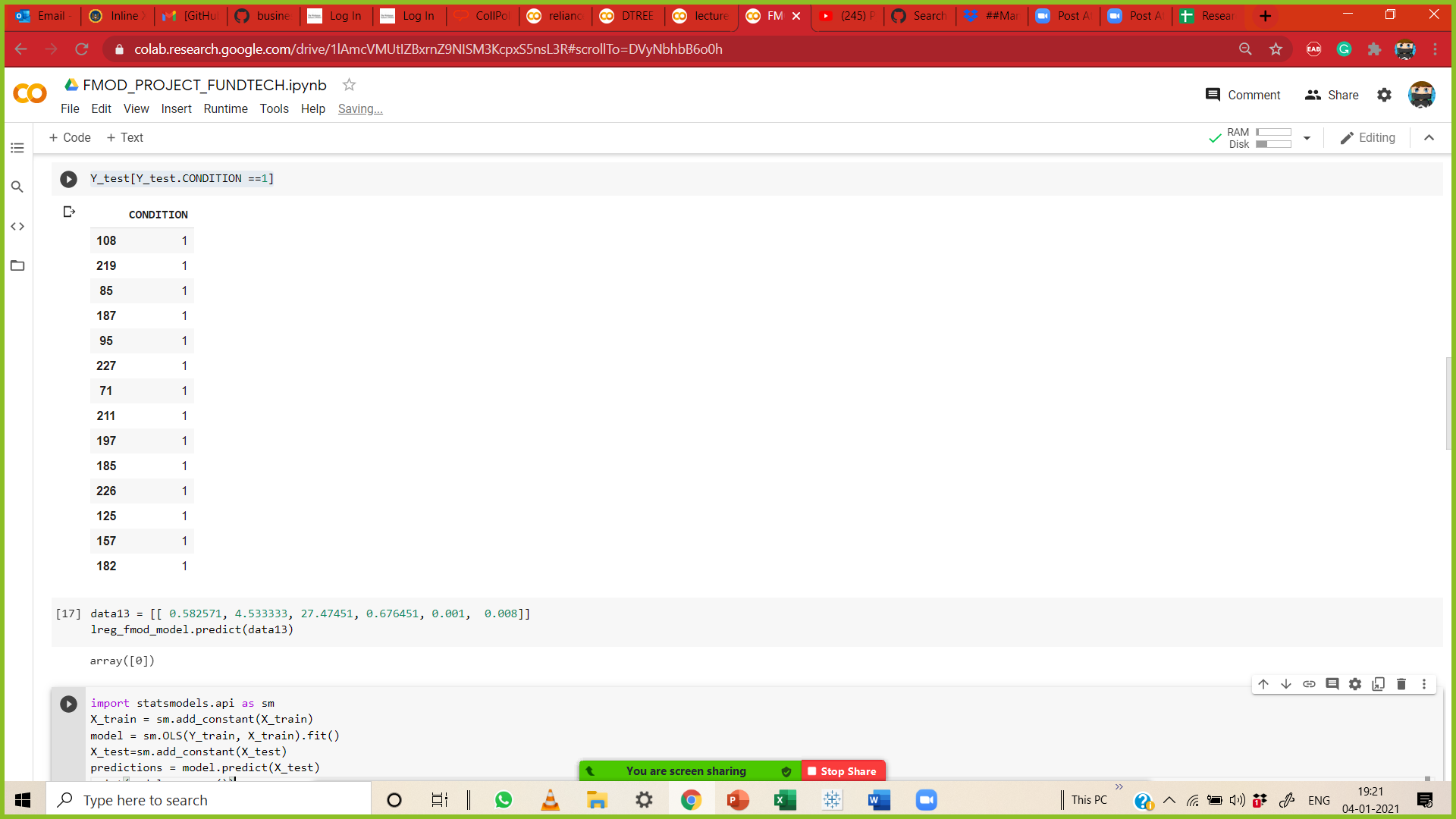
1. **Logistic regression-** Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables.
2. **Decision Tree classifier-** The goal of using a Decision Tree is to create a training model that can use to predict the class or value of the target variable by learning simple decision rules inferred from prior data (training data). In Decision Trees, for predicting a class label for a record we start from the root of the tree.
3. **Random Forest classifier-** The fundamental reason to use a random forest instead of a decision tree is to combine the predictions of many decision trees into a single model. The logic is that a single even made up of many mediocre models will still be better than one good model.

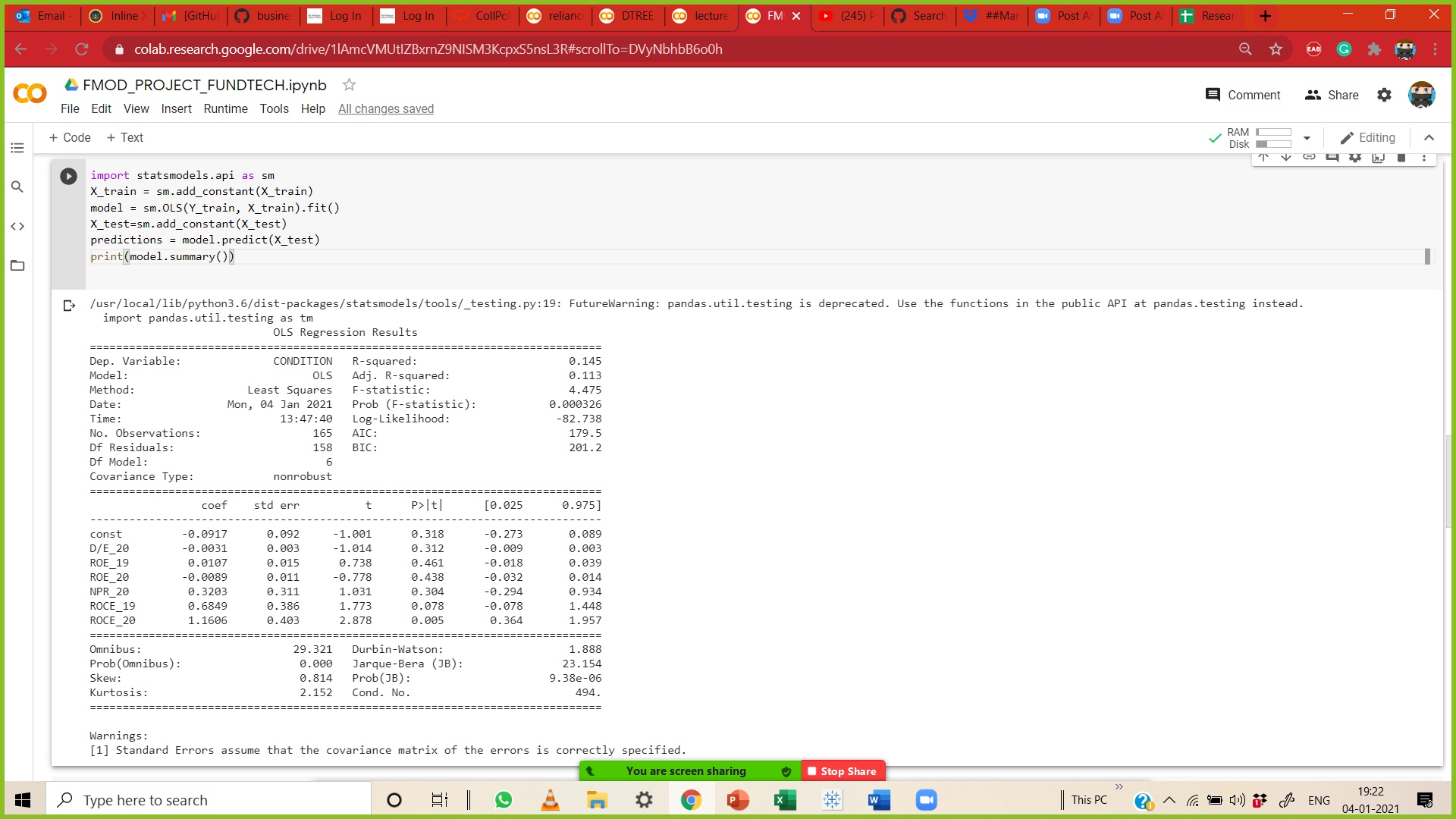
**Model (codes):**

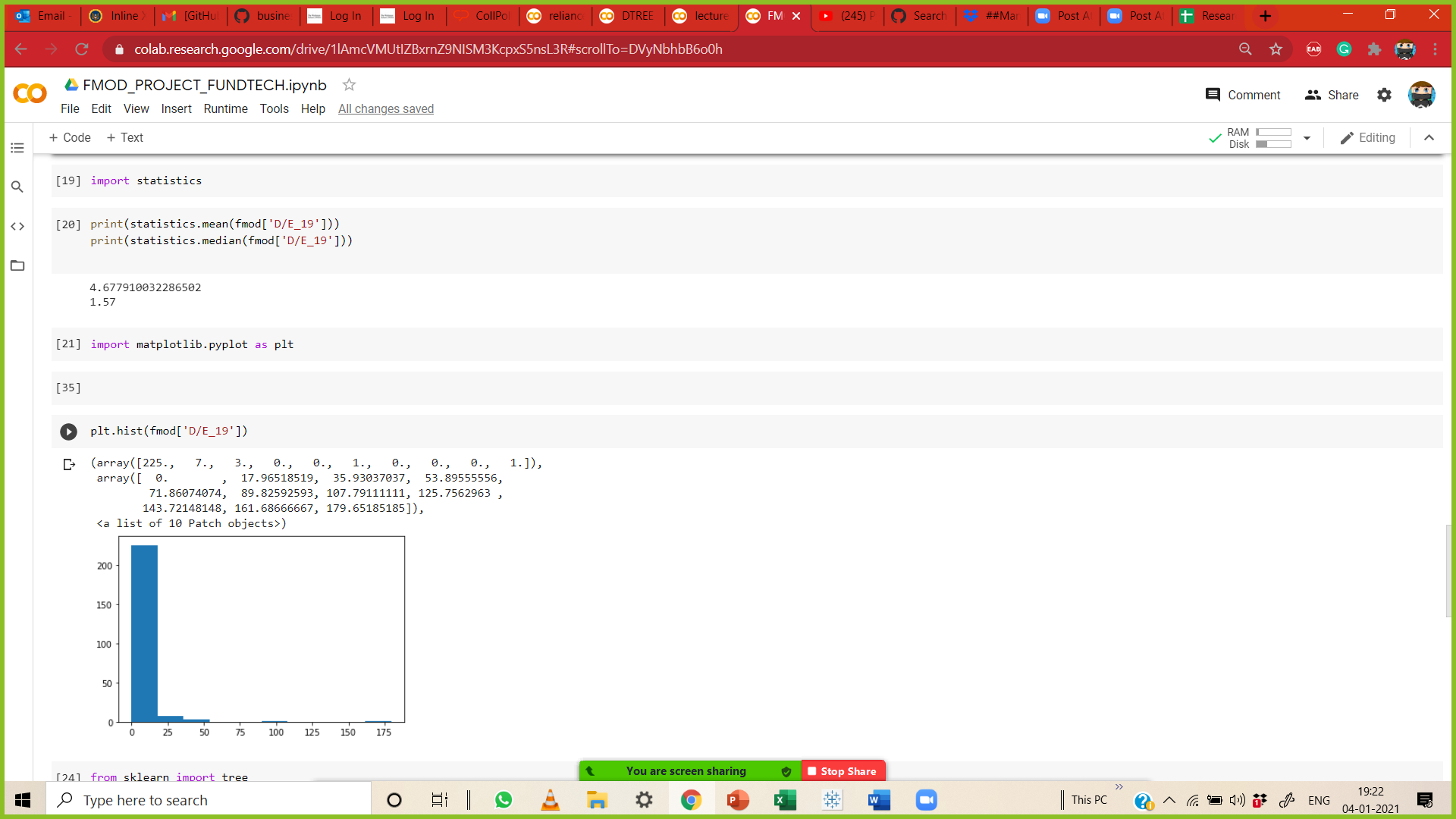


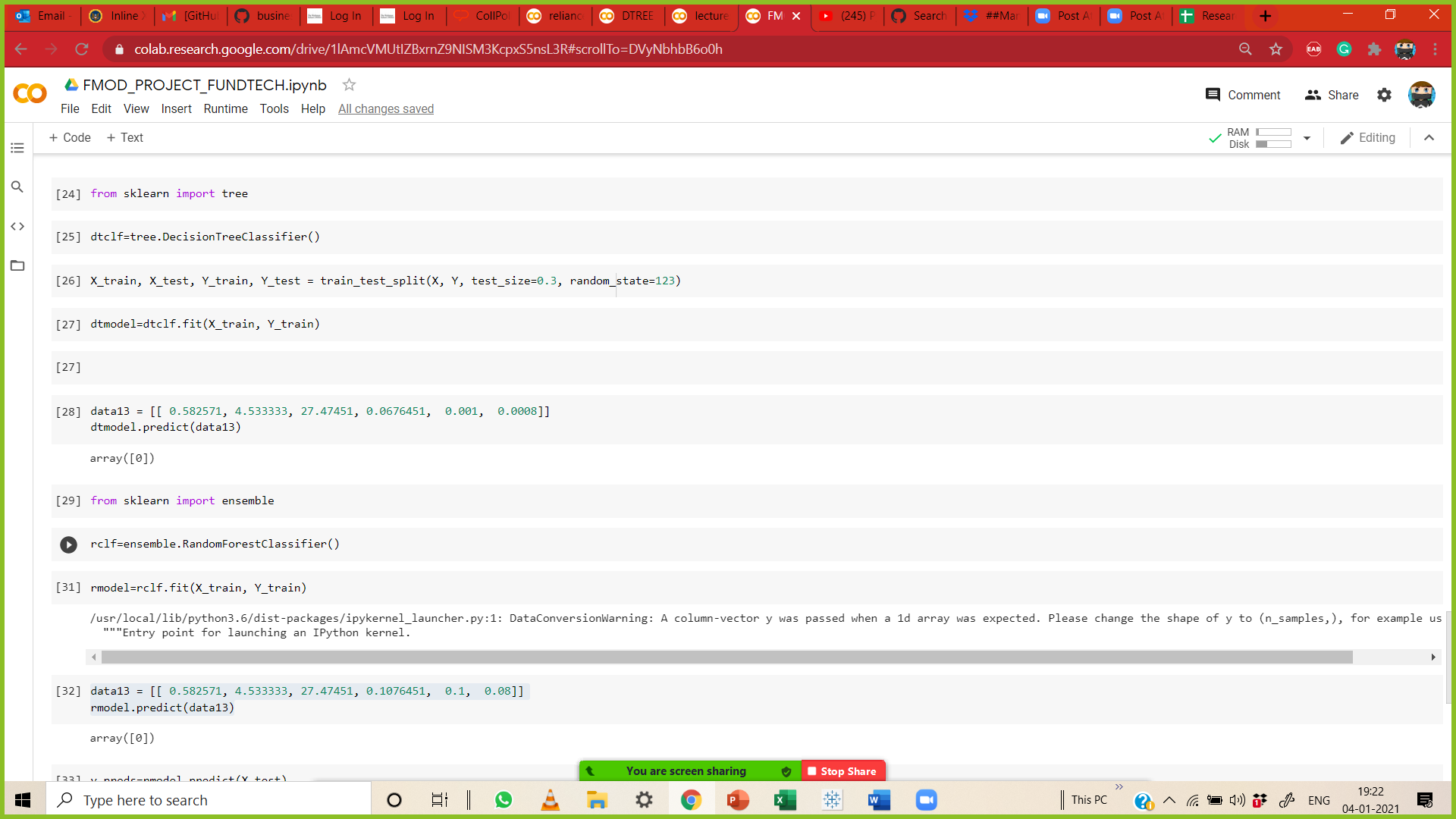


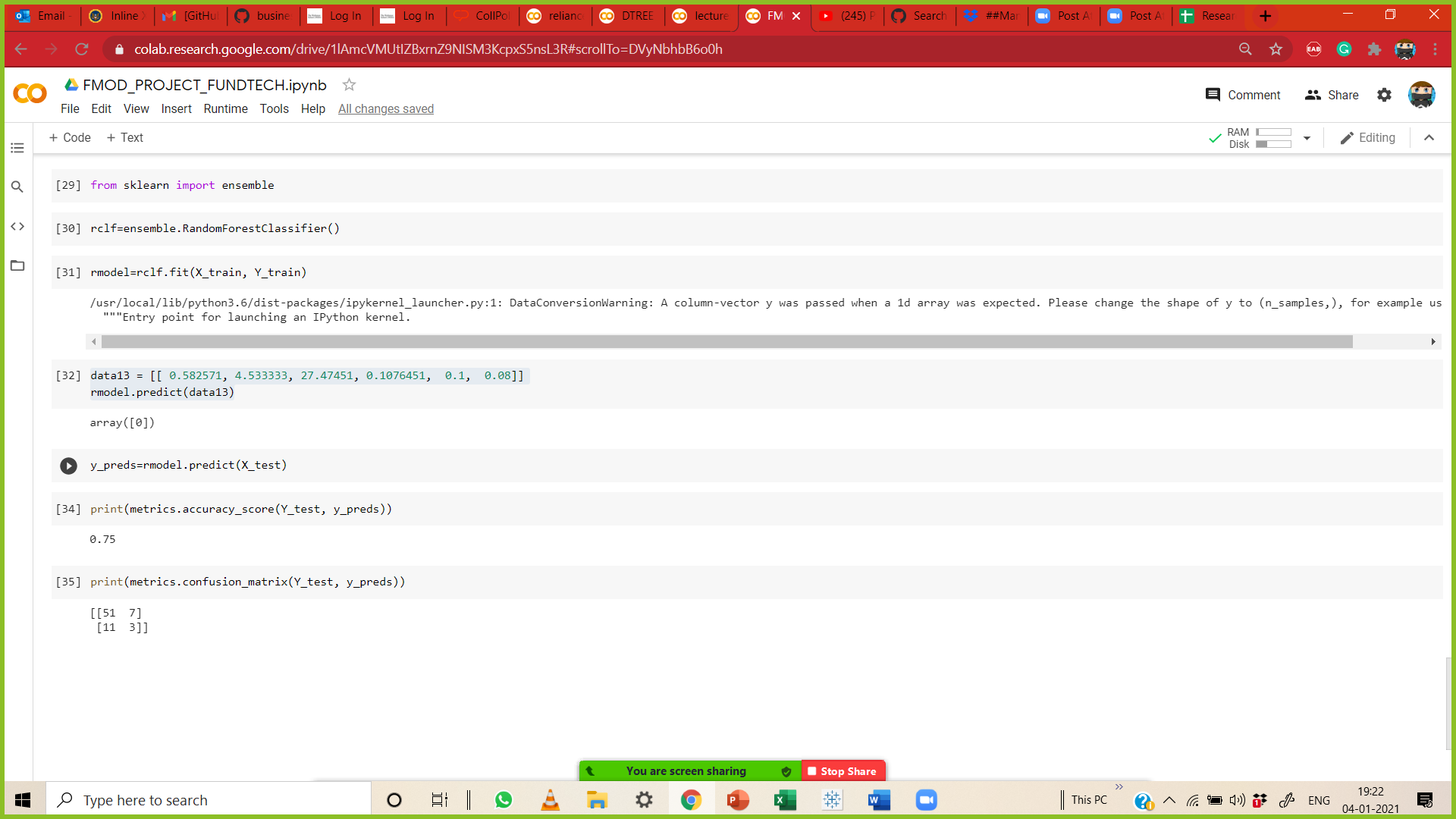












**Prediction and recommendation:**

By using different models, we found that there are few variables which need to remove because of high p value. And finally, we found that our model has 81.99% accuracy score. This model is helpful to identify whether company is financially good or not and based on that banking firm can think for whether to do business with company or not who are seeking loan in the market.

**Summary (statistics and visuals):**

We have selected most relevant and important financial ratio which explains whether company is earning or not and how much they are stable and capable to pay back the stakeholder based on that a banking firm think for working with that company or not. After collecting dataset, we use few models which is relevant for this kind of dataset and split the dataset into training and testing for better optimization. In last we have predicted a random company financial condition with help of the same ratios of that company.

data13 = [[ 3.582571, 4.533333, 27.47451, 0.1076451, 0.1, 0.08]]

rmodel.predict(data13)

array([0])

**In above example following are the sequence of variable-**

|  | **D/E\_ 20,** | **ROE\_ 19,** | **ROE\_ 20,** |  | **NPR\_20,** | **ROCE\_19,** | **ROCE\_20** |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |

And array 0 defines that company’s financial condition is not good. In above case D/E\_20 is 3.58 which is very high as compare to our recommendation i.e. 2 that mean if the D/E is higher than 2 then company has very high debt.

**Recommendation to business:**

**By using this model banking firm can decide whether to proceed with company at initial stage. We are working on this model to improve it and increase their accuracy by finding more suitable variable.**