**About:**

Electoral bonds is a way of funding the political parties. SBI released the data for the year 2019-24. This is a small project in which the data was filtered out from the file and was visualized for basic questions.  
Tech Used:  
1. PySpark

2.Power BI

**Storing the data:**

1. Create database name Analytical.

**Script>>**

Create database if not exists Analytical;

1. We want these table in our database after datasets processing completes.

1.DateOfPurchase(Date)

2.NameOfPurchaser(varchar)

3.PartyName(varchar)

4.Denomination(BigInt)

5.Status(char)(10)

6.Year(char)(4)

**Working with DataSets in CSV formats:**

1. Get the Purchaser Data and Party data in CSV format.
2. Place them in a separate folders.
3. Write a python code for the following tasks.
4. Select only the required columns from both the csv files.

i.Filter the data where status is paid

ii.Format the date of purchase from for example 1-Apr-2019 to 2019-04-01 format.

iii.Create an Year field for further analysis.

1. Now join the both files on the bond number column to get the party names.
2. Now get the data and insert in the database table we have created.

**Create an analytics dashboard:**

1. Top 5 parties that got most of the funds.
2. Top 5 companies that have paid the most
3. Total number of donations per year.
4. Add Slicers and Filter for filter for a particular year

**Dashboard:**

