Data Competition Fall 2020

PPP Loan job retention prediction

Data Dictionary

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Column Type** | **Column Description** |
| **Index** | Int | Loan Identifier |
| **Lender** | String | Organization giving the loan |
| **Loan Range** | String | Amount given as a loan (range) |
| **BusinessType** | String | Categories of businesses (LLC, Trust, Corporation…) |
| **DateApproved** | Object | Dates the loans were approved |
| **NonProfit** | String | Tags for Non-profit organizations. |
| **RaceEthnicity** | String | Ethnic Categories of the organizations taking the loans |
| **Veteran** | String | Indicates if borrower is a Veteran or not |
| **Gender** | String | Gender of the loan receiver |
| **CD** | String | Location of the loan transaction – Congressional District |
| **City** | String | Location of the loan transaction – City Level |
| **State** | String | Location of the loan transaction – State Level |
| **Zip** | Int | Five digits that designated a specific location in the US |
| **NAICSCode** | float | North American Industry Classification System is the standard used by Federal statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data related to the U.S. business economy. |
| **Industry** | String | Business establishments classification |
| **JobsRetained**  **(Target)** | Int | An estimate of previously existing filled positions that are retained as a result of funding by the loan |