Differentially Private Credit Card Clients Report from Taiwan

December 12, 2022

1 Discussion

The database can be found here: https://www.kaggle.com/datasets/uciml/default-of-credit-card-clients-dataset
The github can be found here: https://github.com/nikhilchoppa/cs211-final-project
Our project video can be found here: https://youtu.be/zT8L7tV1GQ8
It should be noted that reported dollar amounts are in the New Tiawan dollar

1.1 Privacy Budget

Overall in the document our epsilon value is 0.1, and from since we invoke the laplace mechanism 33 times, through sequential composition our total privacy budget would be 3.3

2 Statistics

2.1 Basic Averages

The bill_amt and limit_bal differentially private average statistics were generated using the sparse vector technique to determine a clipping parameter for the data, and then generating differentially private sums and counts to find a differentially private average.Whereas the remaining averages and counts were calculated using the good old laplace mechanism or in the case of the averages, we used the method for choosing a clipping which does not require thousands of queries as displayed in Homework 4

2.1.1 Average Age

The average age of all credit card clients is 30.86, the differentially private average age of all customers is 31.21. This gives an error of 1.14%.

2.1.2 Average Credit Limit Balance

The average credit card balance limit of all credit card customers is 167484.32, the differentially private average credit card balance limit of all customers is 70236.44. This gives an error of 58.06%.

2.1.3 Average Bill Amount

The average bill amount of all credit card customers for the month of September is 51223.33, the differentially private average bill amount of all customers is 6747.68. This gives an error of 86.83%.The average bill amount of all credit card customers for the month of August is 49179.08, the differentially private average bill amount of all customers is 40909.14. This gives an error of 16.82%.The average bill amount of all credit card customers for the month of July is 47013.15, the differentially private average bill amount of all customers is 40122.51. This gives an error of 14.66%.The average bill amount of all credit card customers for the month of June is 43262.95, the differentially private average bill amount of all customers is 29700.0. This gives an error of 31.35%.The average bill amount of all credit card customers for the month of May is 40311.4, the differentially private average bill amount of all customers is 30749.95. This gives an error of 23.72%.The average bill amount of all credit card customers for the month of April is 38871.76, the differentially private average bill amount of all customers is 17753.21. This gives an error of 54.33%.

2.1.4 Average Pay Amount

The average pay amount of all credit card customers for the month of September is 5663.58, the differentially private average pay amount of all customers is 5237.27. This gives an error of 7.53%.The average pay amount of all credit card customers for the month of August is 5921.16, the differentially private average pay amount of all customers is 5792.28. This gives an error of 2.18%.The average pay amount of all credit card customers for the month of July is 5225.68, the differentially private average pay amount of all customers is 4416.84. This gives an error of 15.48%.The average pay amount of all credit card customers for the month of June is 4826.08, the differentially private average pay amount of all customers is 4076.45. This gives an error of 15.53%.The average pay amount of all credit card customers for the month of May is 4799.39, the differentially private average pay amount of all customers is 4061.05. This gives an error of 15.38%.The average pay amount of all credit card customers for the month of April is 5215.5, the differentially private average pay amount of all customers is 4770.83. This gives an error of 8.53%.

2.2 Basic Counts

The follwing statistics were generated using the value counts function and applying the laplace mechanism as a lambda function to preserve the dataframe.

2.2.1 Education Levels

The most common education level as determined by using a differentially private method is 2 14025.674934
1 10584.516946
3 4898.217782
5 289.172421
4 119.490531
6 35.926821
0 20.687730
Name: EDUCATION, dtype: float64.

2.3 Conditional Averages

2.3.1 Average Monthly Payments of Male credit card clients

The average monthly credit card payments of male customers is 4939.08 with differential privacy applied.

2.3.2 Average Monthly Payments of Female credit card customers

The average monthly credit card payments of female customers is is 4893.41 with differential privacy applied.

2.3.3 Average Bill Amount credit clients in Higher Education

The average bill amount of all credit card customers in Higher Education is 27613.99 with differential privacy applied.

2.4 Conditional Counts

2.4.1 Most Common Marital Status with defalt 'YES'

The comparison of marital status with if they HAVE defaulted using a differentially private method is MARRIAGE
0 -0.03076
1 3200.96924
2 3335.96924
3 78.96924
Name: 1, dtype: float64.

2.4.2 Most Common Marital Status with defalt 'NO'

The comparison of marital status with if they HAVE NOT defaulted using a differentially private method is MARRIAGE
0 50.555786
1 10454.555786
2 12624.555786

3 240.555786
Name: 0, dtype: float64.