What characterizes sensitive data in your definition?

Here we have the scope of sensitive data outlined by European Commission as of today. Here is a block quote from European Commission’s website. ([European Commission, n.d.](https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_es))

*The following personal data is considered ‘sensitive’ and is subject to specific processing conditions:*

* *personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs;*
* *trade-union membership;*
* *genetic data, biometric data processed solely to identify a human being;*
* *health-related data;*
* *data concerning a person’s sex life or sexual orientation.*

What does privacy mean in automatic data processing?

In the context of automatic data processing, the territory of privacy extends beyond individual personal data stored in a securely protected space (e.g. server, database, etc.) to aggregate statistical summaries published in the public space.

With today’s technologies, even if we do not have data regarding an individual person, aggregate statistics of a dataset, which contains the individual and others, can reveal a significant amount of information about the particular individual contained in the dataset (Vadhan, 2020).

In other words, even if there is no public disclosure about the privacy data at the level of the individual data, summary statistics of the aggregate dataset that contains the individual data can compromise the privacy. (Dwork, 2008, p. 1)

Simply put, privacy in the context of automatic data processing can be compromised without releasing any individual information about a particular person. This is the issue that we need to address.

The theme can be encapsulated into the following one single question: How can we producestatistical information of an aggregate dataset without compromising the privacy of individuals?

What is differential privacy?

Here is a brief characterization of differential privacy by Cynthia Dwork:

“*Roughly speaking, differential privacy ensures that the removal or addition of a single database item does not (substantially) affect the outcome of any analysis. It follows that no risk is incurred by joining the database, providing a mathematically rigorous means of coping with the fact that distributional information may be disclosive.*” (Dwork, 2008, p. 2)

For example, when we remove one row of the observations from the dataset, that changes the aggregate statistics of the modified dataset and the change in the aggregate statistics would inform malicious third parties (like hackers) of the removed row/observation.

In order to make the aggregate statistics of the modified dataset indistinguishable with that of the original dataset, differential privacy introduces a special type of noise into the dataset.

And Laplace distribution provides us with a desirable feature to generate such a noise. (Dwork, 2008, p. 14)

In addition, Professor Natalia Salaberry describes the privacy in the context of automatic data processing in her note in the Google Colab notebook (Salaberry, n.d.). Here is an English translation of her note:

*“Differential Privacy (DP) is a probabilistic mathematical methodology within automatic data processing, which allows organizations to collect, analyze and share aggregate information about individuals, users or customers, while maintaining their privacy.*

*It is about incorporating some distortion on data that needs to be protected, resulting in some noise (Dwork, 2008), but it will not substantially affect a subsequent analysis. If not, it offers the possibility of having a rigorous mathematical method to deal with the possibility of re-identification of the holder. However, it will require the evaluation of a cost, which will be given by how much real information is willing not to be published or used in exchange for obtaining privacy.” (Salaberry, n.d.)*

What is the proposed model?

The proposed model of Differential Privacy attempts to provide “ad omnia guarantee”. In other words, it aims at guaranteeing indistinguishable aggregate statistical properties in the context of automated data processing. And “ad omnia guarantee” can be defined in the following mathematical formula:

Definition 1. A randomized function K gives ε-differential privacy if for all data sets D1 and D2 differing on at most one element, and all S ⊆ Range(K),

Pr[K(D1) ∈ S] ≤ exp(ε) × Pr[K(D2) ∈ S] (1)

Here, “[t]*he mechanism K adds appropriately chosen random noise to”* the dataset. (Dwork, 2008, p. 3) Therefore, K is a random function mechanism of the differential privacy. And ε measures the loss of privacy and the determinant of the robustness of the differential privacy. (Dwork, 2008, p. 2)

ε=0 represents complete privacy. The smaller the ε, the more robust the differential privacy would be.

One essential feature of the *ad omnia guarantee* of differential privacy is thatthe behavior of the mechanism K is independent of any adversary’s knowledge. (Dwork, 2008, p. 3)

In other words, the ad omnia guarantee of Differential Privacy is “*a very strong guarantee, since it is a statistical property about the behavior of the mechanism and therefore is independent of the computational power and auxiliary information available to the adversary/user.*” (Dwork, 2008, pp. 2-3)

These properties extend to group privacy. (Dwork, 2008, p. 3)

Nevertheless, there are some limitations of differential privacy.

First, there is a trade-off between differential privacy and the accuracy. The smaller the ε, the less the accuracy would be. This trade-off can be alleviated by increasing the size of the dataset.

*“For each sample size, as the epsilon value decreases, that is, as the choice is made to reduce the loss of privacy, the error increases. This error represents, in percentage terms, the average variability that exists between the true value in the sample and the value obtained from applying differential privacy. In turn, it is observed that as the sample size increases (last graph), at the same value of the epsilon parameter, the error decreases.”* (Salaberry, n.d.)

Second, differential privacy is not an absolute guarantee of privacy.

“*differential privacy ensures that only a limited amount of additional k is incurred by participating in the socially beneficial databases.*” (Dwork, 2008, p. 3)

Now, the objective is to determine the size of noise that can make statistical property of aggregate dataset insensitive to change in a row of the dataset.

In order to determine how much noise should be incorporated into the data in order to balance the loss of privacy, the sensitivity of the weight of an individual's data in the calculations made must be measured (Dwork, 2008).

Where the query is a function f, and the database is X, the true answer is the value f(X), ***sensitivity*** is defined as follows.

Definition 2. For f :D →, the sensitivity of f is

for all differing in at most one element.

“*Note that sensitivity is a property of the function alone, and is independent of the database. The sensitivity essentially captures how great a difference (between the value of f on two databases differing in a single element) must be hidden by the additive noise generated by the curator.*”

The scaled symmetric exponential distribution with standard deviation , denoted , has mass at x proportional to . … On query function *f* the privacy mechanism ***K*** responds with *f*(X) + adding noise with distribution independently to each of the *k* components of *f*(X).

Theorem 1. For *f* : D → , the mechanism that adds independently generated noise with distribution to each of the k output terms enjoys ε-differential privacy. (Dwork, 2008, p. 5)

Perform a differential privacy model assessment on any variable that might be considered sensitive in a dataset you have or have used.

Obtuve el conjunto de datos predeterminado del [sitio web de R-Data](https://r-data.pmagunia.com/dataset/r-dataset-package-islr-default#:~:text=The%20Default%20data%20set%20is,into%20a%20variable%20called%20Default.). (Games, Witten, Hastie, & Tibshirani, 2017)

The dataset contains 4 variables: default, student, income, and bank account balance and has 10,000 observations.

I found the information about income and bank account balance relevant to sensitive data.

I performed the two types of analyses.

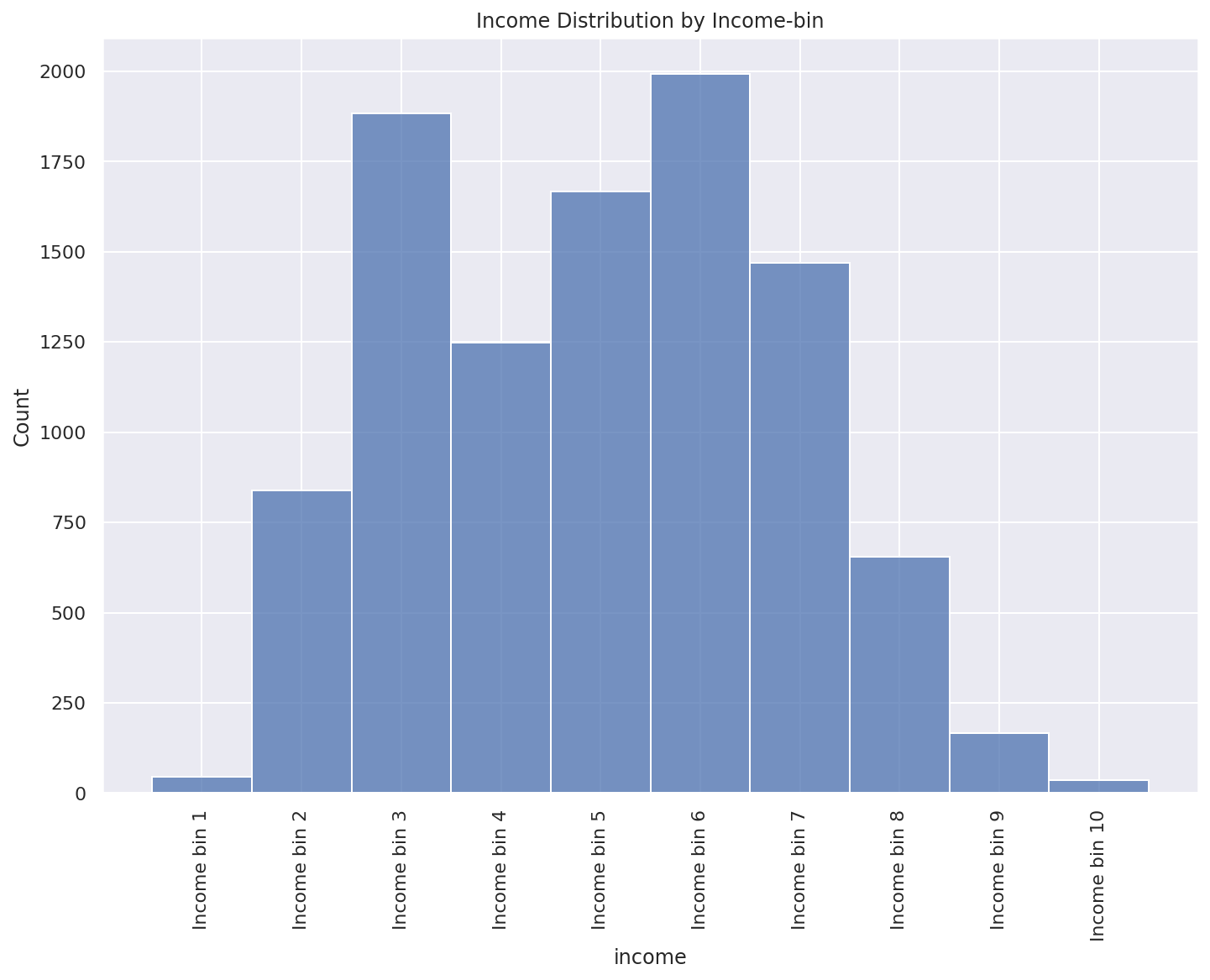
* Laplace Mechanism's impact on the dataset:
* Laplace Mechanism's impact on the model results of Logistic Regression

ails of these analyses are provided in the notebook of my python code in the link here.

1. Laplace Mechanism's impact on the dataset

I applied the Laplace mechanism of differential privacy to ‘income’ data and replicated the exercises performed by the notebook provided by Professor Natalia Salaberry.

Since ‘income’ is a continuous numeric data, I allocated the ‘income’ data into 10 bins of categorical data.

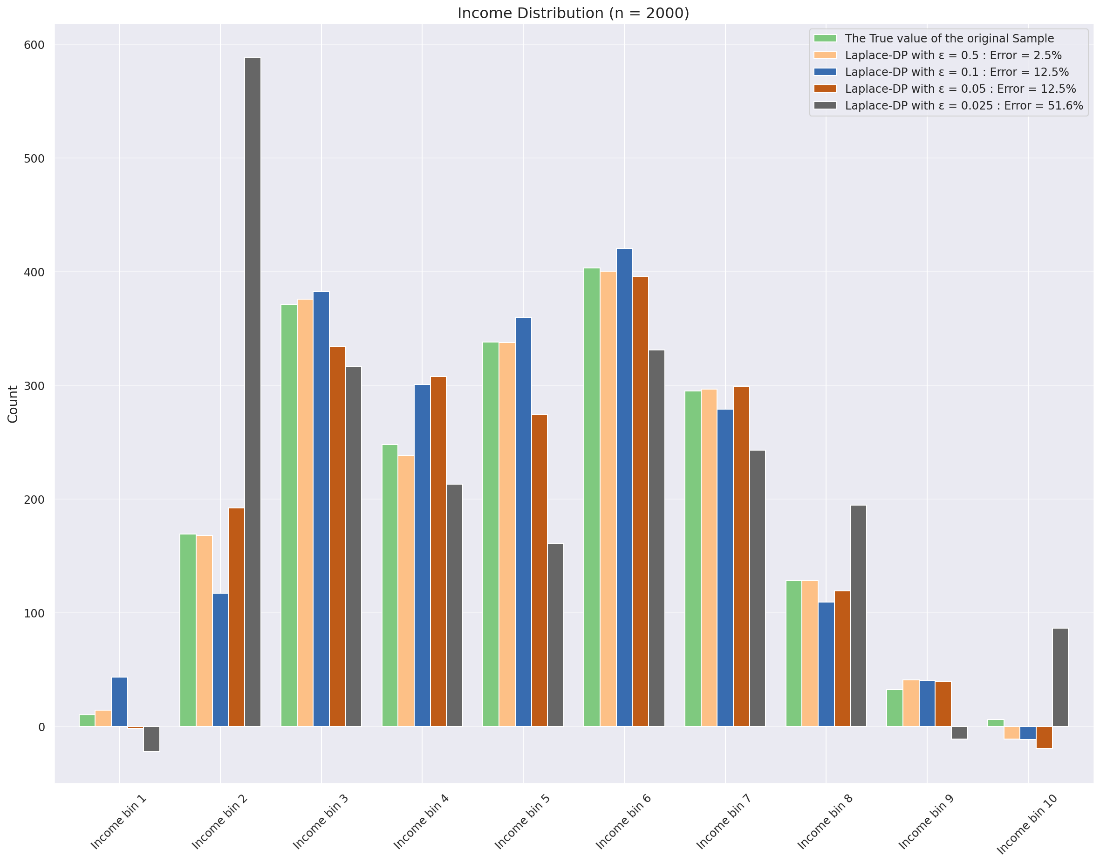


Here, I will present the results of histograms for 4 different sample sizes: 2,000, 3,000, 5,000, and 8,000.

One interesting finding that we had never encountered in the example demonstrated by Professor Salaberry is that for this case Laplace mechanism generated a negative count for the two ends of the spectrum: “income bin 1” and “income bin 10”. some of ‘income’ bins. Although I sought a remedy for this problem, I could not find any viable solution. We need to impose some constraint on the result of the Laplace noise in order to maintain the relevancy of the dataset to the real life situation. This needs to be addressed in the future.

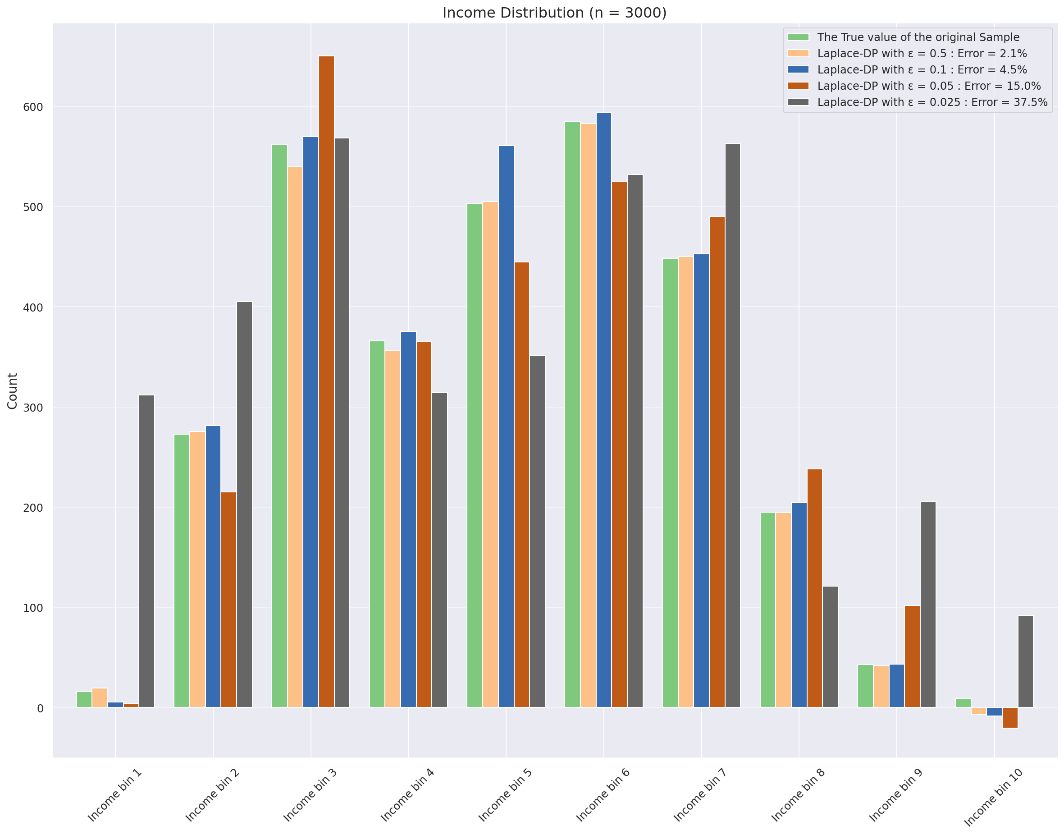
1. Sample Size of 2,000

Here, we can identify negative results on the both ends of the spectrum: “income bin 1” and “income bin10”.



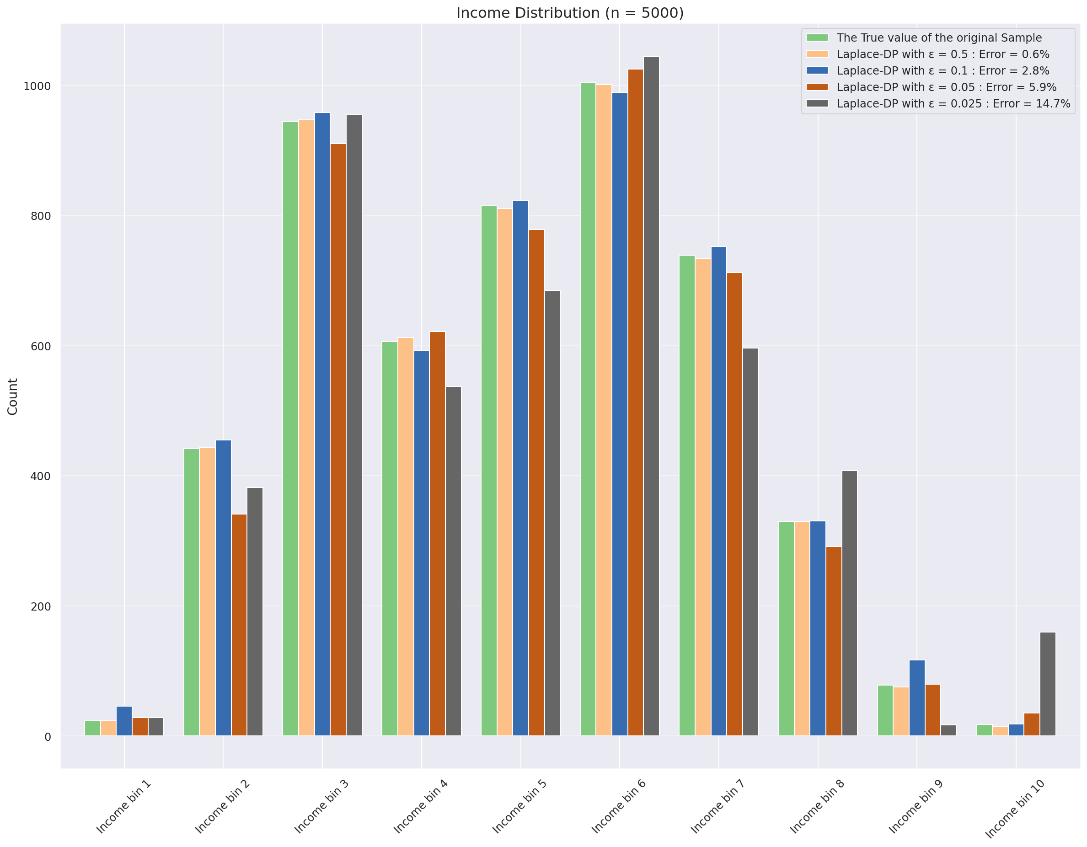
1. Sample Size of 3,000

In this example, we can identify negative results in “income bin10”.



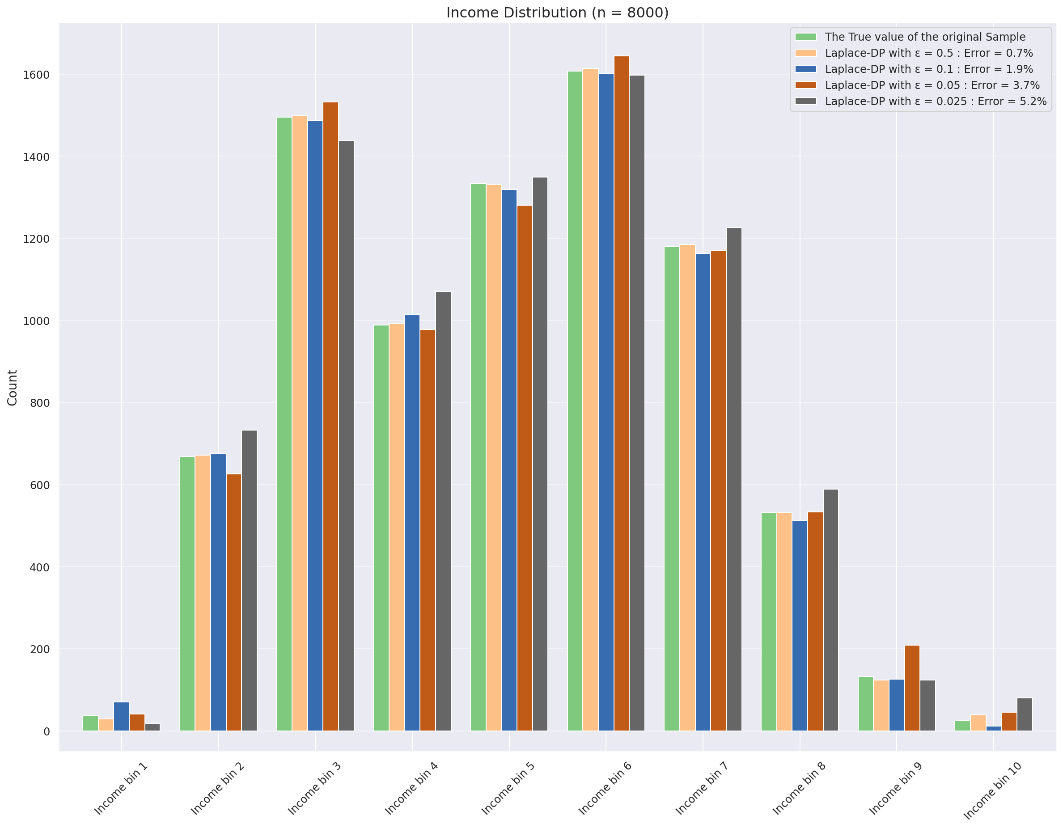
1. Sample Size of 5,000

For this sample size of 5,000, 50% of the total population, we have no more negative cases.



1. Sample Size of 8,000

The distributions of data with noise from Laplace mechanism of differential privacy.



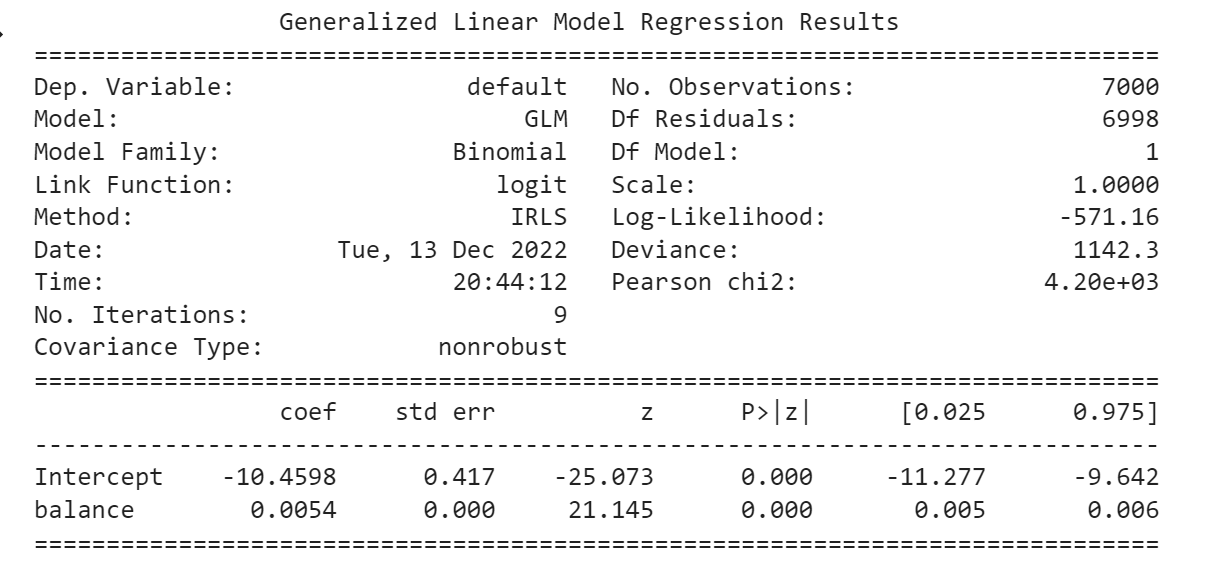
1. Laplace Mechanism's impact on the model results of Logistic Regression

I applied the Laplace mechanism of differential privacy to ‘balance’ data, instead of ‘income’, to see whether the noise from the Laplace mechanism will make a material change on the result of Logistic Regression.

Here, we can compare the regression results of the logistic regression on these two datasets: original and the one with noise from Laplace mechanism.

* Original Dataset:

Here is the summary of the regression result on the original dataset.



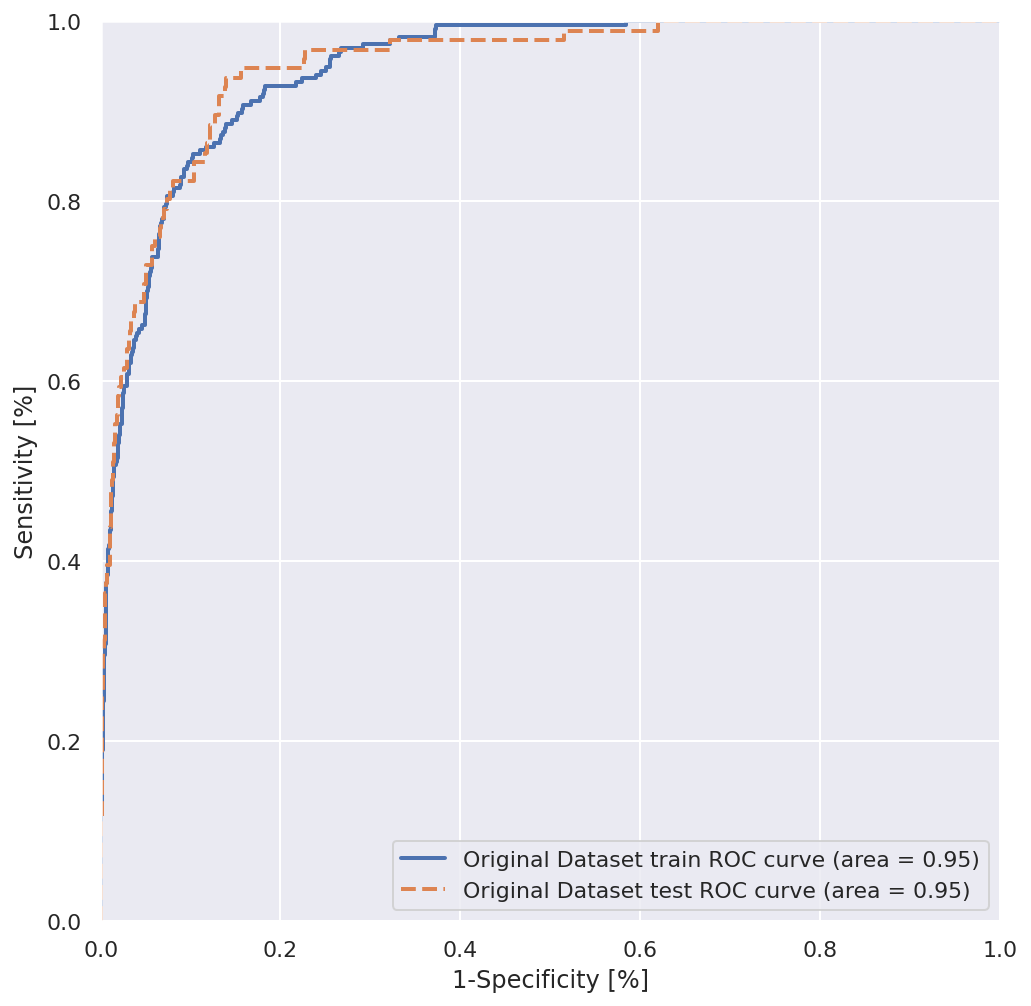
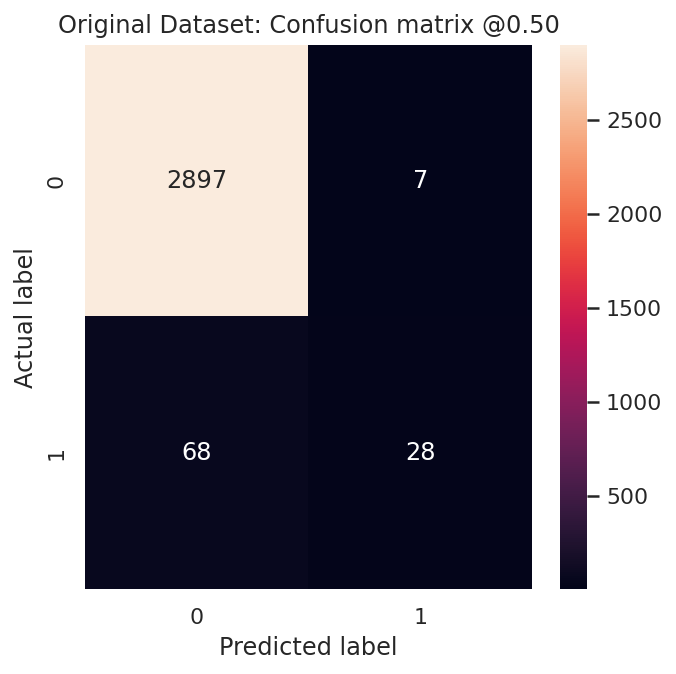
The intercept has its z-score at a large negative value (-25.073) with its negligible size of p-value. The coefficient of ‘balance’ has its z-score at a large positive value (21.145) with its negligible size of p-value. Thus, there is no particular reason to reject these results.

The fitted regression model is as follows:

Probability[default] =

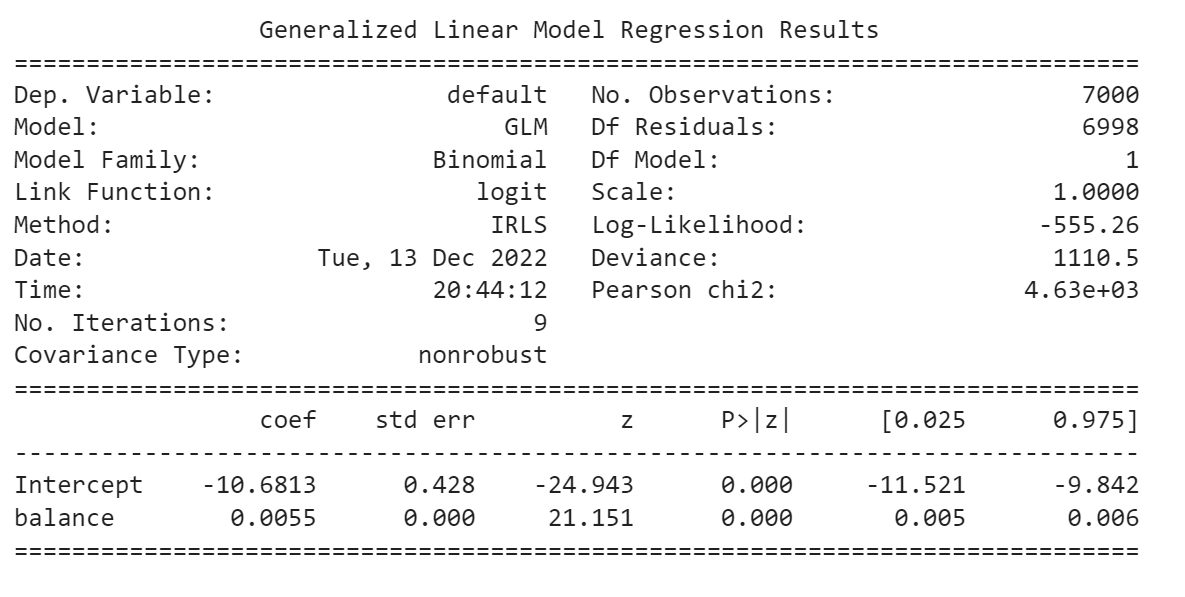
Here, is a sigmoid function (logistic function).

Here are the Confusion Matrix and the ROC Curve of the model on the original dataset.



* Laplace Dataset:

Now we can examine the regression model on the dataset with noise from Laplace mechanism of differential privacy.



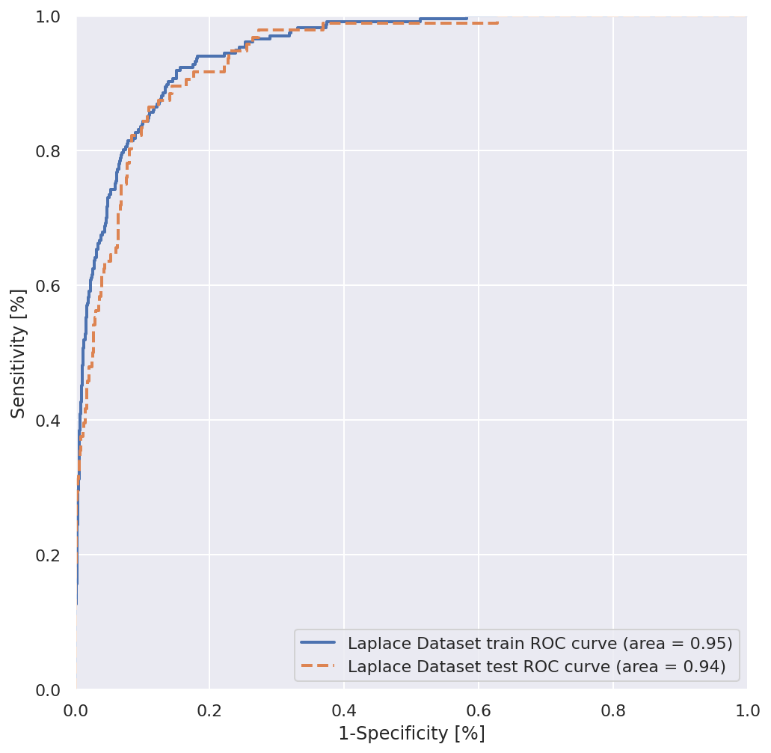
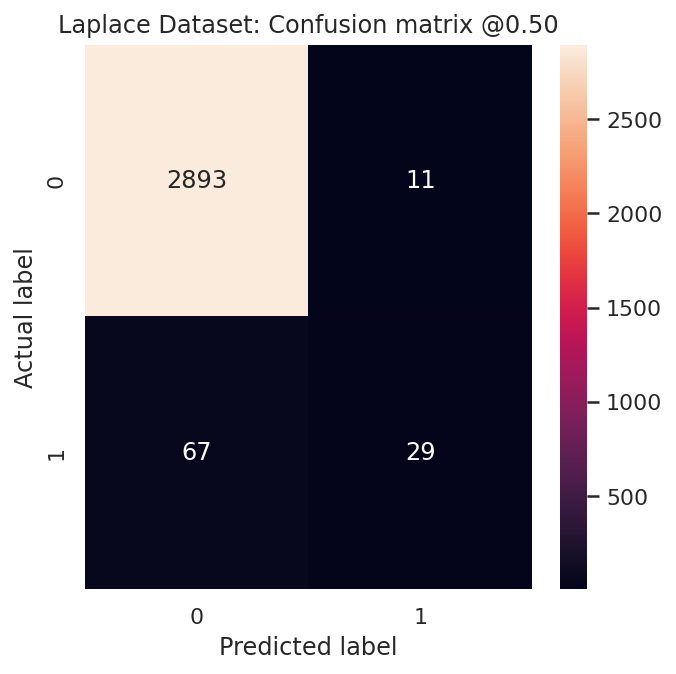
The intercept has its z-score at a large negative value (-24.943) with its negligible size of p-value. The coefficient of ‘balance’ has its z-score at a large positive value (21.151) with its negligible size of p-value. Thus, there is no particular reason to reject these results.

The fitted regression model is as follows:

Probability[default] =

Here, is a sigmoid function (logistic function).

Furthermore, we can observe the Confusion Matrix and the ROC Curve on the dataset with noise from Laplace mechanism of differential privacy.



Now, comparing all these results, we do not observe material differences between the regression results on these two datasets.

Overall, we could conclude that Laplace mechanism would not materially affect the model result.

Explain why you consider it sensitive

Both ‘income’ and ‘balance’ are financial information. Financial information is sensitive because it describes the social status of a person.

Income can be used to cluster the salary potential of workers. And the information about the current income of a person could determine the person’s cluster. In other words, it can be used to trap the person in a particular income cluster and limit the future job potential of the person.

Hackers can misuse the information about bank account balance to target people with a particular financial profile to wage cyber-attacks.

Financial information should be treated as sensitive in these contexts.

Reference

Dwork, C. (2008, 4). *Differential Privacy A Survey of Results.* Retrieved from Microsoft: https://www.microsoft.com/en-us/research/wp-content/uploads/2008/04/dwork\_tamc.pdf

European Commission. (n.d.). *What personal data is considered sensitive?* Retrieved from European Commission: https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive\_en#:~:text=personal%20data%20revealing%20racial%20or,sex%20life%20or%20

Salaberry, N. (n.d.). *Introducción a la Privacidad Diferencial con Python.* Retrieved from Google Colab: https://colab.research.google.com/drive/1l-84B6F94HjFz1BJj-5D-qUBlvmTbScW?usp=sharing

Vadhan, S. (2020, 5 16). *The Science Behind WhiteNoise: Differential Privacy.* Retrieved from YouTube: https://youtu.be/TMMHL-6ezkg

# 文献目録

Dwork, C. (2008, 4). *Differential Privacy A Survey of Results.* Retrieved from Microsoft: https://www.microsoft.com/en-us/research/wp-content/uploads/2008/04/dwork\_tamc.pdf

European Commission. (n.d.). *What personal data is considered sensitive?* Retrieved from European Commission: https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive\_en#:~:text=personal%20data%20revealing%20racial%20or,sex%20life%20or%20

Salaberry, N. (n.d.). *Introducción a la Privacidad Diferencial con Python.* Retrieved from Google Colab: https://colab.research.google.com/drive/1l-84B6F94HjFz1BJj-5D-qUBlvmTbScW?usp=sharing

Vadhan, S. (2020, 5 16). *The Science Behind WhiteNoise: Differential Privacy.* Retrieved from YouTube: https://youtu.be/TMMHL-6ezkg