**Part 4**

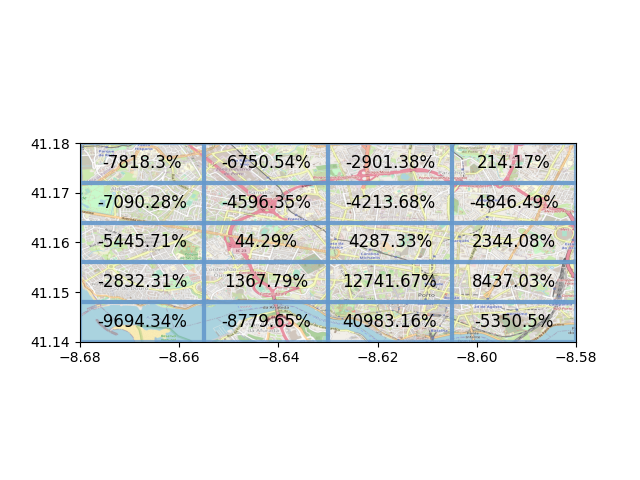
**Experimental Analysis**

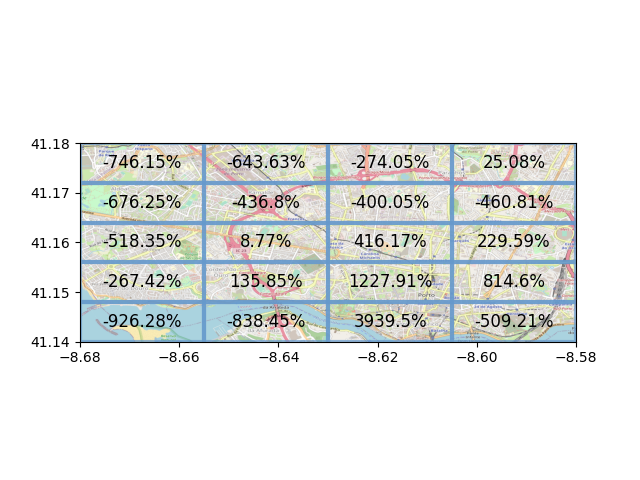
**Error Analysis**

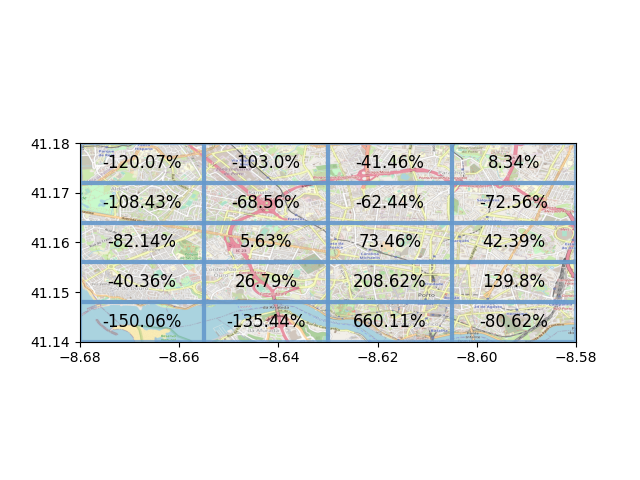
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Protocol/Epsilon** | **0.01** | **0.1** | **0.5** | **1** | **2** |
| GRR | 128,065 | 9,229 | 1,477 | 0,739 | 0,211 |
| RAPPOR | 38,287 | 6,833 | 4,979 | 5,122 | 4,994 |
| OUE | 28.336,983 | 2.694,457 | 429,607 | 161,463 | 45,952 |

The above table contains resulting errors for epsilon values when a local differential privacy protocol used. As it can be seen from the table the error decreases as the epsilon values increases, implying lower privacy for higher amount of epsilons. However, Rappor causes an outlier behavior when epsilon equals to 1 and error increases and when epsilon equals 2 error decreases again, which makes Rappor non-monotonic. We need higher error rates for higher privacy, so by looking at the table we can say that OUE has a very high amount of error for every epsilon value which makes it best protocol for higher privacy. However, very high privacy rates may cause utility losses, so the protocol must be chosen carefully for the tasks. On the other hand, the fastest protocol is GRR and the slowest protocol is OUE. Since we need to behave different for 0 bits and 1 bits, calculating perturbation probabilities takes more time than the other protocols.

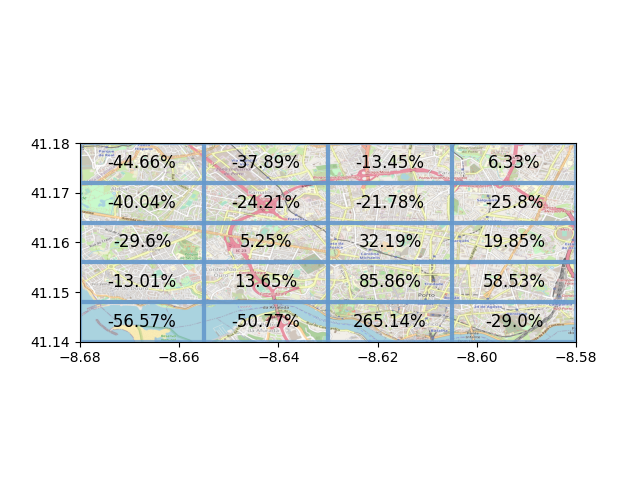
**Visual Analysis**

Epsilon = 0.01

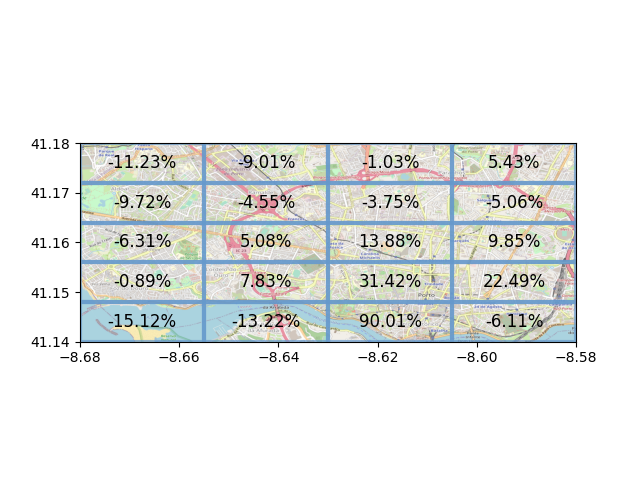
Epsilon = 0.1

Epsilon = 0.5

Epsilon = 1



Epsilon = 2



The figures above are resulting taxi locations in Porto after GRR protocol applied to the data for epsilon values 0.01, 0.1, 0.5, 1 and 2. As we can see from the figures as epsilon value increases the percentages become more realistic. In the map for epsilon 0.01, we observe lots of small percentages such as -746.15% and big percentages such as 1227.91%, which not realistic. However, this confusion provides a high privacy by hiding real taxi percentage from the end user. As the epsilon increases, the percentages get in a more realistic boundary. For example, when we set epsilon to 2, we see bigger negative percentages as -11.23% and smaller positive percentages as 31.42%. By inspecting those map figures, we can conclude that smaller epsilon values are better for hiding information from the end users and adversaries. However, high levels of privacy can cause utility loss. Therefore, we should determine our aim and set the epsilon value accordingly. If we need high levels of privacy, we should use small epsilon values.