**COMP 4651 Cloud Computing and Big Data Systems - Project Report**

# Analysis of Incident Reports from San Francisco Police Department

**Group members:**

*Jakob KALLESTAD, Guillaume JOUET-PASTRE, Tom HAZET*

**Abstract:**

(What we have done, the aim of the project and short about the results we found)

**Background:**

(A little introduction to our dataset, San Francisco police department, their method of collecting data)

**Technical details:**

(We did a little experimenting… then we deployed a ec2 cluster which we ran a databricks notebook on… we stored the dataset on amazon s3 (my bucket name is databricksbucketjakob))

**Results:**

(We found out the following:

* Categories of crimes likely to lead to an arrest —> ???
* How categories of crimes are correlated to location —> ???
* Show correlation between season/month/weekday/hour —> ???

)

**Conclusion:**

(blablabla)

**Appendix:**

(<https://github.com/GuillaumeJouetp/crime_insights.git>)