

Team C Work Breakdown Report

Team C: Min Chen x19205040, Jia Lin x22117644

Team C	Min Chen	Jia Lin
Data Source	Dataset 1: https://catalog.data.gov/dataset/covid-19-community-vulnerability-crosswalk-rank-ordered-by-score	Dataset 2: https://rapidapi.com/axisbits-axisbits-default/api/covid-19-statistics/
<i>Project report:</i>		
Abstract	Abstract for the project and dataset 1	Abstract for the project and dataset 2
Introduction	Introduction for the project and dataset 1	Introduction for the project and dataset 2
Related Work	Related work for dataset 1	Related work for dataset 2
Methodology	Dataset 1: Methodology flow Data acquiring, Database management Python and ML algorithms Jupyter in Anaconda, PostgresDB Data processing and modelling, Decision tree analysis, Data visualization Datasets combination	Dataset 2: ETL pipeline Data acquiring Database management Programming language Data processing Data Visualisation Datasets combination
Result Evaluation	Result and evaluation for dataset 1 and combination of two datasets	Result and evaluation for dataset 2 and combination of two datasets
Conclusion & Future Work	Conclusion and future work for dataset 1 and the project	Conclusion and future work for dataset 2 and the project
Acknowledgement & References	Acknowledgement for this project and references for dataset 1	Acknowledgement for this project and references for dataset 2
<i>Code artefact:</i>		
	Dataset1 file overview Dataset1 processing and ML Analysis Dataset1 more analysis with Datasets combination	DBAP_CA2_teamC_JiaLin_Extract_Transform_1_Global_Data.ipynb DBAP_CA2_teamC_JiaLin_Extract_Transform_2_US_Data.ipynb DBAP_CA2_teamC_JiaLin_Extract_Transform_3_US_by_State.ipynb DBAP_CA2_teamC_JiaLin_Load_PostgreSQL_Visualisation_Combination.ipynb
<i>Video presentation:</i>		
	1-6 Slides	7-12 Slides