

Data And Information Quality Project

- **PROJECT ID:** 21
- **PROJECT NUMBER:** 1
- **ASSIGNED DATASET:** USERS
- **STUDENT:** PASQUALE CASTIGLIONE 10657816
- **ASSIGNED TASK:** CLUSTERING

Contents

1	SETUP CHOICES	3
1.1	Chosen ML algorithms	3
1.2	Chosen ML performance evaluation metrics	3
1.3	Imputation techniques selected	3
2	PIPELINE IMPLEMENTATION	3
2.1	Description of the steps you performed	3
3	RESULTS	3
3.1	Description of the main results obtained	3
3.2	ML performance comparison between the imputation/outlier detection techniques you have implemented	3

1 SETUP CHOICES

1.1 Chosen ML algorithms

1.2 Chosen ML performance evaluation metrics

1.3 Imputation techniques selected

2 PIPELINE IMPLEMENTATION

2.1 Description of the steps you performed

3 RESULTS

3.1 Description of the main results obtained

3.2 ML performance comparison between the imputation/outlier detection techniques you have implemented