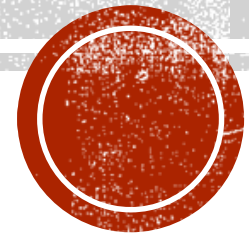


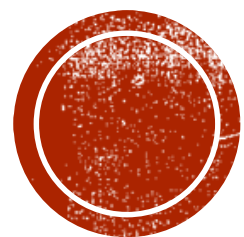
IBM CAPSTONE PROJECT: MACHINE LEARNING TO IDENTIFY AND MAP SCRUM ADOPTION



OVERVIEW

1. Definition
2. Problem
3. Methodology
4. Results
5. Discussion
6. Conclusion



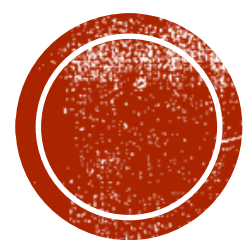


DEFINTION



- Scrum is an Agile software development framework. It is the most adopted and successful Agile framework used globally.
- Statistics taken from VersionOne.com indicate in the region of 90 percent of organizations that use an Agile methodology or framework implement Scrum as part of their solution.
- Is it possible to identify different states or cities that displays successful and unsuccessful Scrum adoption?



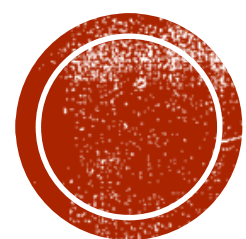


PROBLEM



- Our problem will identify South Africa's (SA) Agile organizations as the population sample.
- Our investigation will gather and analyze data from these organizations within the provinces in SA to explore the mapping of the Scrum adoption within the country.
- Our solution is to use the data gathered from an online survey on organizations that currently use Scrum for their daily projects.





METHODOLOGY



- Data Sources

- The data was gathered using an online survey questionnaire.

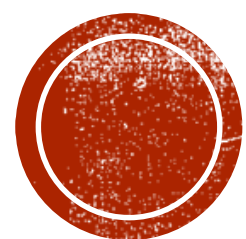
- Data Cleaning

- Data cleaning was minimal. The majority of the data cleaning that occurred was in the Jupyter notebook.

- Data Analysis

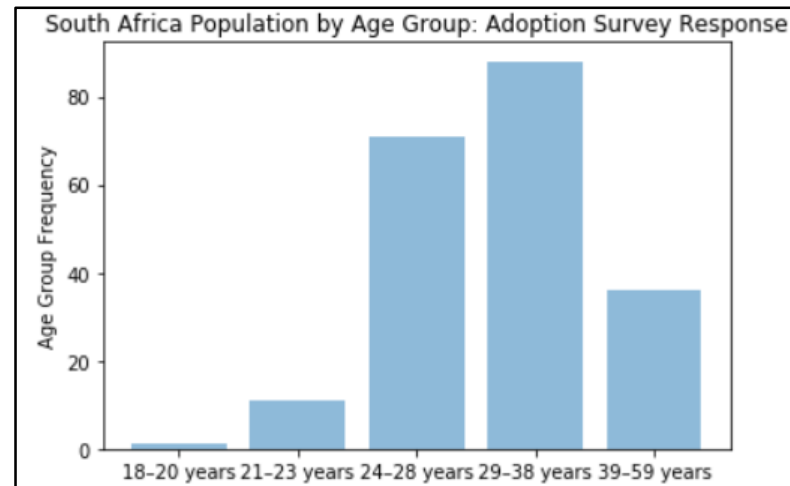
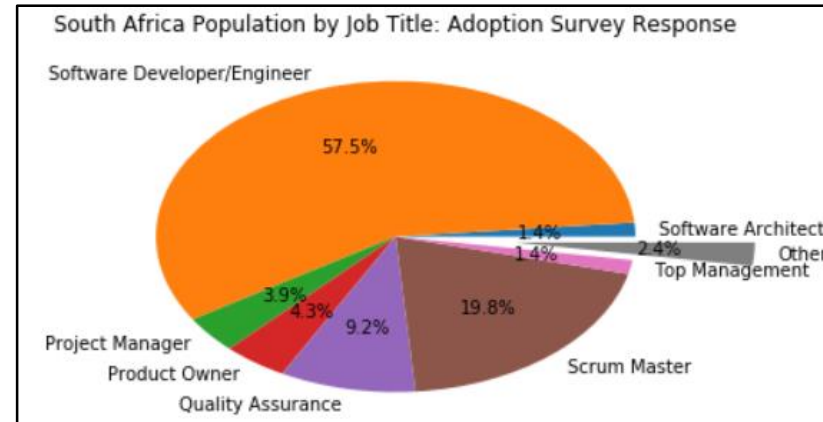
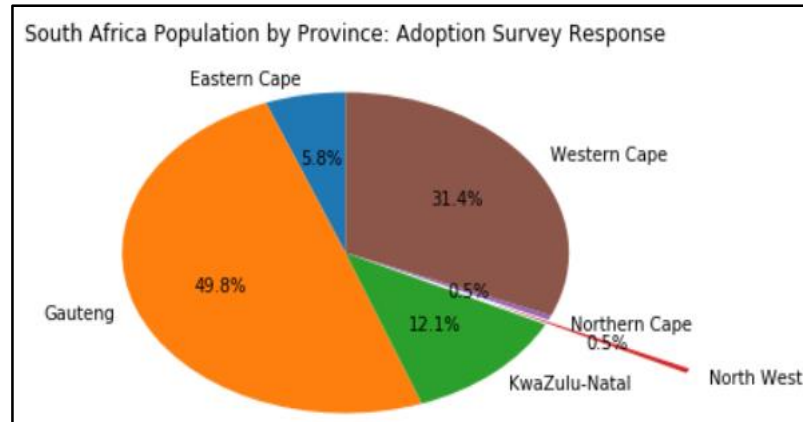
- The analysis was broken down into 3 major sections.
 - Descriptive analysis
 - EFA and MLR model building
 - Mapping of statistical results





RESULTS

- Demographics: The demographics of the survey respondents have been broken up into Province, Job Title, and Age Group.

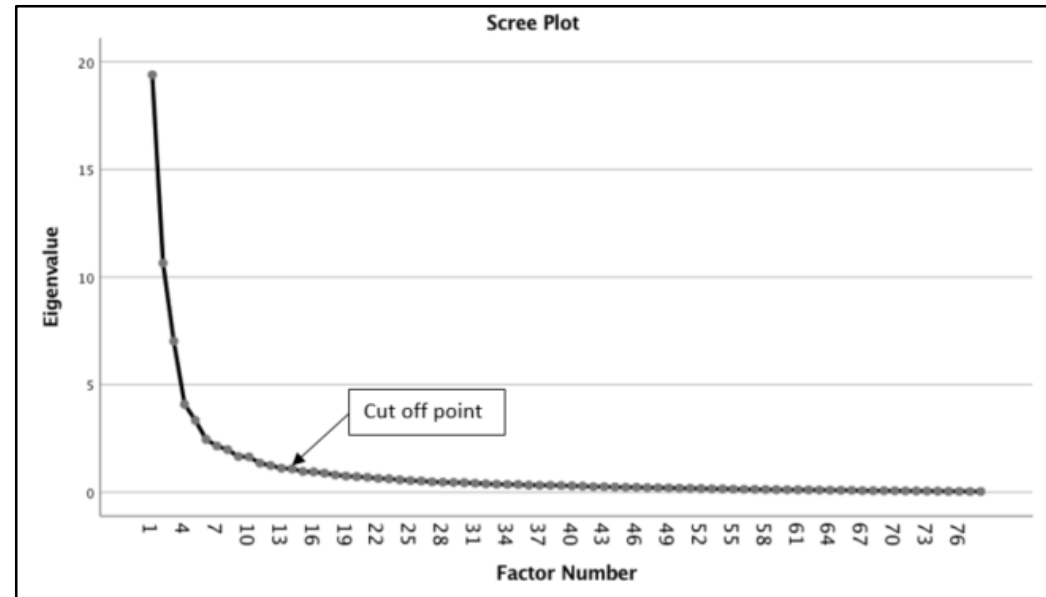


■ Descriptive statistics: The basic descriptive stats

	WorkWithinSA_A1	ScrumUsage_A2	AgeGroup_A3	Province_B1	JobTitle_B2	WorkExperience_B3	WorkWithinOrg_B4
count	207.0	207.000000	207.000000	207.000000	207.000000	207.000000	207.000000
mean	1.0	4.420290	4.710145	4.932367	5.594203	4.743961	3.125604
std	0.0	1.132989	0.832031	2.856651	14.832383	1.313540	1.485748
min	1.0	2.000000	2.000000	1.000000	1.000000	1.000000	1.000000
25%	1.0	4.000000	4.000000	3.000000	2.000000	4.000000	2.000000
50%	1.0	5.000000	5.000000	3.000000	2.000000	5.000000	3.000000
75%	1.0	5.000000	5.000000	9.000000	5.000000	6.000000	4.000000
max	1.0	7.000000	6.000000	9.000000	99.000000	7.000000	7.000000



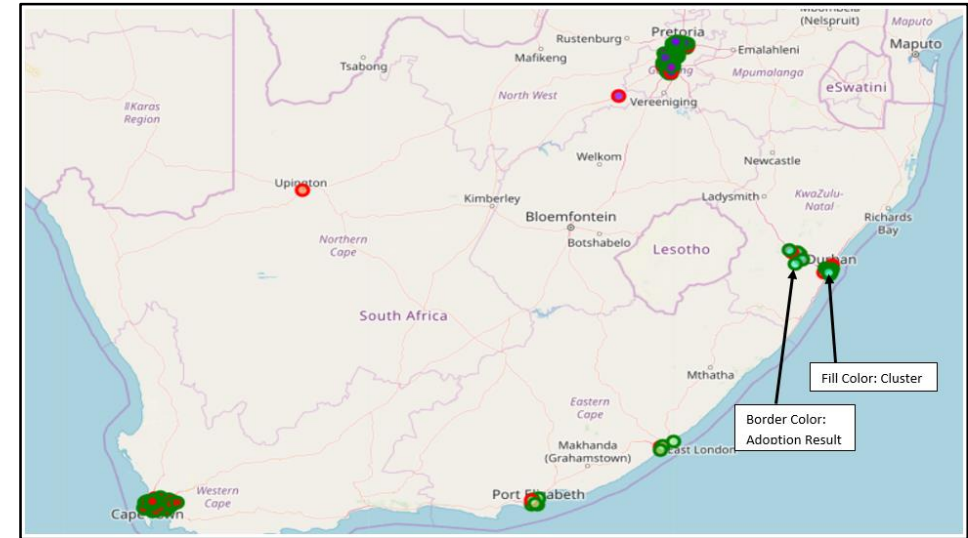
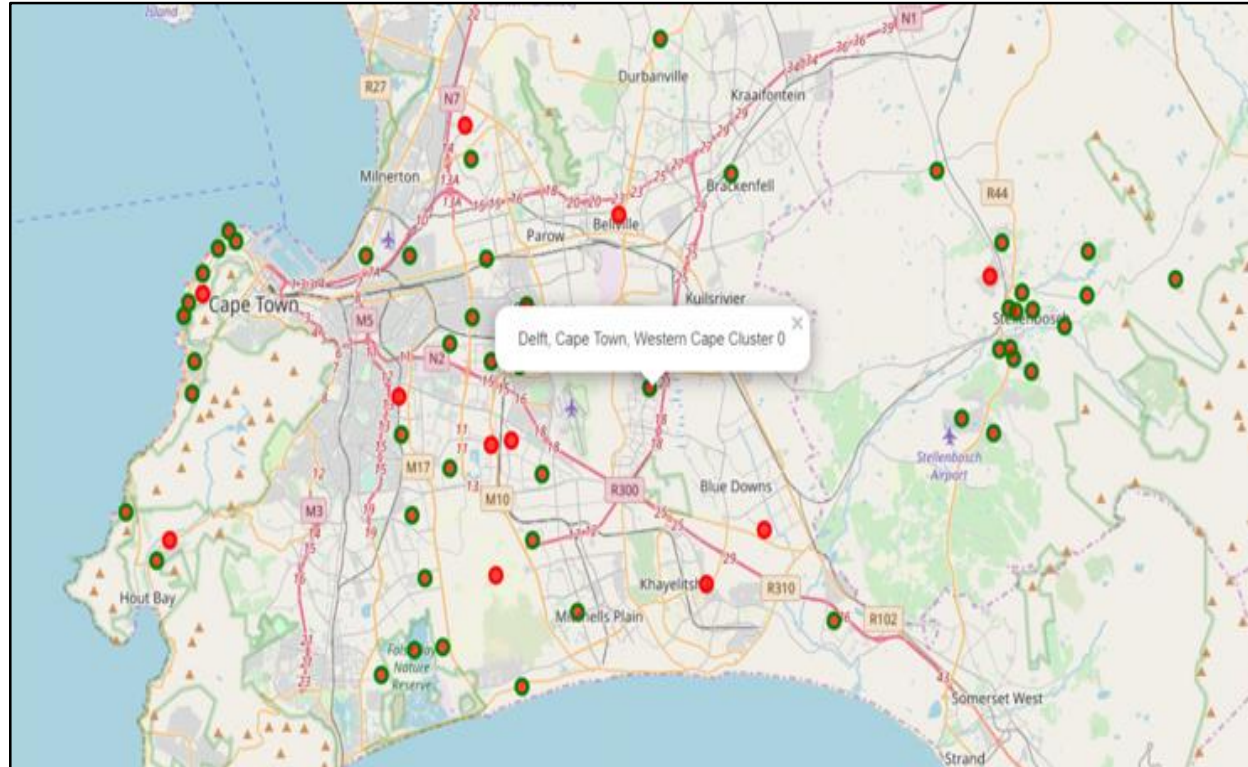
- EFA and MLR model: The modelling and statistical analysis



```
Coefficients: [[ 0.2629548 -0.1718043 -0.04751236 -0.19088155 -0.02120981 -0.01685441
 0.03589423 -0.00990481 0.07244454 0.08612357 -0.2187448 0.12017564
 0.06509107 -0.10111133 0.21034699 0.03975154 -0.0890368 -0.12790865
 0.10027975 0.01105817 -0.02143907 0.15537124 -0.11637622 -0.11546516
 -0.05987753 0.10319157 -0.11755821 0.06521124 -0.04029331 0.16922373
 -0.18814219 -0.06201037 -0.19195809 0.17046082 0.3264677 -0.33218474
 -0.05486219 -0.05714516 0.05818345 0.19909266 -0.12107272 0.06207647
 -0.15043556 -0.03228923 -0.05517595 0.08404214 0.17129823 -0.21190744
 0.01057313 -0.00549803 -0.10445267 0.15882121 0.06909275 0.20975819
 -0.01171534 -0.32892371 0.30253181 0.05731448 -0.10796836 -0.1716132
 0.12520852 0.16267298 -0.35043984 0.00587423 -0.13991979 -0.09221308
 0.26572039 -0.00343887 0.32798381 0.03161179 -0.22868844 0.0543215
 0.2932309 0.39826162 -0.0269742 -0.09684351 0.10509685 0.10696488]]
```

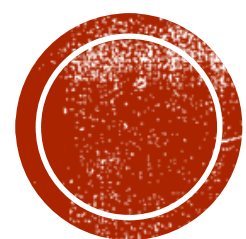


- The Adoption Mapping: The analysis and generation of the adoption plotting and mapping based on province clustering.



	Adoption	Rejection
ProvinceMappingToB1		
Eastern Cape	10	2
Gauteng	88	15
KwaZulu-Natal	17	8
North West	0	1
Northern Cape	0	1
Western Cape	53	12





DISCUSSION



- The findings from the results indicates that Scrum adoption rates is successful with more than 81 percent of organizations adopting Scrum within South Africa (SA).
- The machine learning model generates variable accuracy between ~0.40% and ~0.80%.
- The demographics of the population sample signifies the interest of the greater SA community in the Agile adoption. Of the 9 provinces in SA, 6 of them captured responses.





CONCLUSION

- This project therefore looked at doing a quantitative study to identify the adoption rates and mapping it onto the SA landscape.
- The limitations of the study is the need for a larger population sample and the need to conduct additional Exploratory Factor Analysis and logistic regression to refine the analysis.
- The author would also like to further refine the model with the assistance of a much larger dataset.

