# Syracuse University

# MS Applied Data Science

# Portfolio

# By: Brandon Croarkin

# SUID: 839788036

# Email: [bcroarki@syr.edu](mailto:bcroarki@syr.edu)

# GitHub Link: <https://github.com/bcroarkin2726/SyracusePortfolio>

# GitHub Table of Contents

1. **2008ArrivalDelayAnalysis**
   1. *Croarkin\_Group2\_Presentation.pptx* – PowerPoint presentation given to class for final project.
   2. *Croarkin\_Group2\_ProjectSummary.docx* – written paper on the 2008 Arrival Delay Analysis.
   3. *FinalProjectDescription.pdf* – overview document of Final Project requirements.
   4. *IST687\_FinalProjectAnalysis.Rmd* – script used to complete project analysis.
   5. *README* – overview of the project including team members and a link to the data used (too large to store in GitHub)
   6. *VisualizingFlightData2.rmd* – script used to create visuals of airplane flights
   7. *carrierAA.pdf* – example output from the VisualizingFlightData2 script of the flight patters of American Airlines over the US.
2. **EuropeanSoccerMatchAnalysis**
   1. *Croarkin\_FinalProject.docx* – written report of analysis submitted for final project.
   2. *Croarkin\_FinalProjectPresentation.pptx* – PowerPoint slides presented to class for final project.
   3. *EPLMatchPrediction.rmd* – R Markdown file of code used to perform analysis.
   4. *Match3.csv* – intermediary file used to create Team\_Record\_Formatted.csv.
   5. *README* – overview of the project, including team members.
   6. *Team\_Record.xlsx* - intermediary file used to create Team\_Record\_Formatted.csv.
   7. *Team\_Record\_Formatted.csv* – formatted data file made in Excel from Match3.csv and Team\_Record.xlsxused to complete analysis.
3. **ShelterAnimalOutcomeAnalysis**
   1. *AustinEmployment.xlxs* – Austin employment data pulled from Bureau of Labor Statistics (BLS) to complement data.
   2. *Brandon\_Croarkin\_Final\_Project\_Code.py* – python script used to perform project analysis.
   3. *Brandon\_Croarkin\_Final\_Project\_Report.docx* – written report of shelter animal outcomes submitted for final project.
   4. *IST652 – FinalProject.ipynb* – JupyterNotebook version of python script used to perform project analysis.
   5. *README* – overview of the project, including team members.
   6. *Shelter Animal Outcomes.pptx* – PowerPoint slides presented to class for final project
   7. *test.csv* – test data pulled from <https://www.kaggle.com/c/shelter-animal-outcomes> to assess performance of models created.
   8. *train.csv* – train data pulled from <https://www.kaggle.com/c/shelter-animal-outcomes> to create models.
4. **SpamEmailClassification**
   1. *1208\_NLP\_Final\_Project.ipynb* – JupyterNotebook python script used to create project analysis.
   2. *Croarkin\_Mak\_Yeap\_FinalProjectReport.pdf* – written report of spam email classification analysis submitted for final project.
   3. *README* – overview of the project, including team members.
   4. *Final\_project.doc* – overview of the final project, including description of the dataset and requirements for the project.
   5. *Spam.zip* – zipped folder containing two sub-folders (Spam and Ham) that respectively contain the emails used for the analysis.
5. **Data Science Resume.pdf** – current professional resume
6. **DataScience@Syracuse Milestone Porfolio Overview Final.pdf** – overview document of Porfolio
7. **OverviewDocument.docx** ­– overview document providing student identification (name, SUID number, email, GitHub link) and table of contents of the GitHub repository.
8. **README.md** – overview of the contents of the GitHub repository.
9. **WrittenPaper.docx** – written paper that contains each learning goal of the program.