**Portfolio Report**

M.S. Applied Data Science Program Reflection

Michael Johnson

February 12, 2022

SUID: 516408416

GitHub: https://github.com/AirJohnson3/Portfolio\_Project

Text

Description automatically generated

Table of Contents

[Introduction 2](#_Toc95579238)

[Course Highlight 1: Data Warehouse and Database Administration Concepts and Database Management 3](#_Toc95579239)

[Course Description 4](#_Toc95579240)

[Learning Objectives 4](#_Toc95579241)

[Project Requirements 4](#_Toc95579242)

[Project Development 4](#_Toc95579243)

[Project Reflection 4](#_Toc95579244)

[Course Highlight 2: Natural Language Processing 3](#_Toc95579245)

[Course Description 4](#_Toc95579246)

[Learning Objectives 4](#_Toc95579247)

[Project Requirements 4](#_Toc95579248)

[Project Development 4](#_Toc95579249)

[Project Reflection 4](#_Toc95579250)

[Course Highlight 3: Big Data 3](#_Toc95579251)

[Course Description 4](#_Toc95579252)

[Learning Objectives 4](#_Toc95579253)

[Project Requirements 4](#_Toc95579254)

[Project Development 4](#_Toc95579255)

[Project Reflection 4](#_Toc95579256)

[Course Highlight 4: Applied Machine Learning for Data Analysis 3](#_Toc95579257)

[Course Description 3](#_Toc95579258)

[Learning Objectives 3](#_Toc95579259)

[Project Requirements 3](#_Toc95579260)

[Project Development 4](#_Toc95579261)

[Project Reflection 4](#_Toc95579262)

[Conclusion 5](#_Toc95579263)

[References 6](#_Toc95579264)

# Introduction

Data collection and subsequent analysis maintained a significant role throughout history and remains vital to the decision-making process today. The science behind understanding that data drives some of the most important advancements through human history, from collecting the positions of stars to fighting the COVID-19 pandemic through information analysis. Companies and governments throughout the world rely on some level of descriptive, predictive, or prescriptive analytics all driven by leveraging data science to derive insights into problems or questions.

The base level of all data starts with a problem or question that could range from a basic “How can we make more money?” to advanced problems like developing self-driving cars. The “science” aspect in data science elicits the idea that undertaking data science problems follows the similar scientific method as many other scientific fields, which all start with asking a question. The initial question or questions of interest drive the acquisition or collection of data for understanding and then provide insight into the modeling and deployment processes as part of the complete data science lifecycle (The Team Data Science Process lifecycle, 2022).

Understanding and working within the scope of the data science lifecycle is an important part of analyzing and building machine learning models. That knowledge provides the foundation of data science within the Applied Data Science graduate program at Syracuse University. Through the program, students identified actionable insights with data analytics through building visualizations and predictive models relying on a mixture of theory and application (Applied Data Science Master’s Degree, 2022). Leveraging an understanding for the tools like Python, R programming, and Excel for analyzing data and building models was an integral part of succeeding in the program.

The following portfolio serves to exemplify the seven learning objectives critical in the Applied Data Science program (Stinnett, 2022):

1. Describe a broad overview of the major practice areas in data science.
2. Collect and organize data.
3. Identify patterns in data via visualization, statistical analysis, and data mining.
4. Develop alternative strategies based on the data.
5. Develop a plan of action to implement the business decisions derived from the analyses.
6. Demonstrate communication skills regarding data and its analysis for relevant professionals in their organization.
7. Synthesize the ethical dimensions of data science practice.

# Course Highlight 1: Data Warehouse and Database Administration Concepts and Management

## Course Description

Data collection and subsequent analysis maintained a significant role throughout history and remains vital to the decision-making process today. The science behind understanding that data drives some of the most important advancements through human history, from collecting the positions of stars to fighting the COVID-19 pandemic through information analysis. Companies and governments throughout the world rely on some level of descriptive, predictive, or prescriptive analytics all driven by leveraging data science to derive insights into problems or questions.

The base level of all data starts with a problem or question that could range from a basic “How can we make more money?” to advanced problems like developing self-driving cars. The “science” aspect in data science elicits the idea that undertaking data science problems follows the similar scientific method as many other scientific fields, which all start with asking a question. The initial question or questions of interest drive the acquisition or collection of data for understanding and then provide insight into the modeling and deployment processes as part of the complete data science lifecycle (The Team Data Science Process lifecycle, 2022).

## Learning Objectives

After taking this course, the students will be able to:

* Describe fundamental data and database concepts
* Explain and use the database development lifecycle
* Create databases and database objects using popular database management system products
* Solve problems by constructing database queries using Structured Query
* Language (SQL)
* Design databases using data modeling and data normalization techniques
* Develop insights into future data management tool and technique trends Recommend and justify strategies for managing data security, privacy, audit/control, fraud detection, backup and recovery
* Critique the effectiveness of DBMS in computer information systems

## Project Requirements

INSERT PROJECT INSTRUCTIONS

## Project Development

INSERT PROJECT

## Project Reflection

From a data engineering standpoint, the data warehouse and database management objectives help in efficiently capturing, organizing, and promoting adequate data storage. Data consisting of text, pictures, or any other data type and can form as structured, semi-structured, or unstructured and requires knowledge for effectively managing the flow in and out of enterprise systems.

Data collection and subsequent analysis maintained a significant role throughout history and remains vital to the decision-making process today. The science behind understanding that data drives some of the most important advancements through human history, from collecting the positions of stars to fighting the COVID-19 pandemic through information analysis. Companies and governments throughout the world rely on some level of descriptive, predictive, or prescriptive analytics all driven by leveraging data science to derive insights into problems or questions.

The base level of all data starts with a problem or question that could range from a basic “How can we make more money?” to advanced problems like developing self-driving cars. The “science” aspect in data science elicits the idea that undertaking data science problems follows the similar scientific method as many other scientific fields, which all start with asking a question. The initial question or questions of interest drive the acquisition or collection of data for understanding and then provide insight into the modeling and deployment processes as part of the complete data science lifecycle (The Team Data Science Process lifecycle, 2022).

# Course Highlight 2: Natural Language Processing

## Course Description

Data collection and subsequent analysis maintained a significant role throughout history and remains vital to the decision-making process today. The science behind understanding that data drives some of the most important advancements through human history, from collecting the positions of stars to fighting the COVID-19 pandemic through information analysis. Companies and governments throughout the world rely on some level of descriptive, predictive, or prescriptive analytics all driven by leveraging data science to derive insights into problems or questions.

The base level of all data starts with a problem or question that could range from a basic “How can we make more money?” to advanced problems like developing self-driving cars. The “science” aspect in data science elicits the idea that undertaking data science problems follows the similar scientific method as many other scientific fields, which all start with asking a question. The initial question or questions of interest drive the acquisition or collection of data for understanding and then provide insight into the modeling and deployment processes as part of the complete data science lifecycle (The Team Data Science Process lifecycle, 2022).

## Learning Objectives

After taking this course, the students will be able to:

* Describe fundamental data and database concepts
* Explain and use the database development lifecycle
* Create databases and database objects using popular database management system products
* Solve problems by constructing database queries using Structured Query
* Language (SQL)
* Design databases using data modeling and data normalization techniques
* Develop insights into future data management tool and technique trends Recommend and justify strategies for managing data security, privacy, audit/control, fraud detection, backup and recovery
* Critique the effectiveness of DBMS in computer information systems

## Project Requirements

INSERT PROJECT INSTRUCTIONS

## Project Development

INSERT PROJECT

## Project Reflection

Data collection and subsequent analysis maintained a significant role throughout history and remains vital to the decision-making process today. The science behind understanding that data drives some of the most important advancements through human history, from collecting the positions of stars to fighting the COVID-19 pandemic through information analysis. Companies and governments throughout the world rely on some level of descriptive, predictive, or prescriptive analytics all driven by leveraging data science to derive insights into problems or questions.

The base level of all data starts with a problem or question that could range from a basic “How can we make more money?” to advanced problems like developing self-driving cars. The “science” aspect in data science elicits the idea that undertaking data science problems follows the similar scientific method as many other scientific fields, which all start with asking a question. The initial question or questions of interest drive the acquisition or collection of data for understanding and then provide insight into the modeling and deployment processes as part of the complete data science lifecycle (The Team Data Science Process lifecycle, 2022).

# Course Highlight 3: Big Data

## Course Description

Data collection and subsequent analysis maintained a significant role throughout history and remains vital to the decision-making process today. The science behind understanding that data drives some of the most important advancements through human history, from collecting the positions of stars to fighting the COVID-19 pandemic through information analysis. Companies and governments throughout the world rely on some level of descriptive, predictive, or prescriptive analytics all driven by leveraging data science to derive insights into problems or questions.

The base level of all data starts with a problem or question that could range from a basic “How can we make more money?” to advanced problems like developing self-driving cars. The “science” aspect in data science elicits the idea that undertaking data science problems follows the similar scientific method as many other scientific fields, which all start with asking a question. The initial question or questions of interest drive the acquisition or collection of data for understanding and then provide insight into the modeling and deployment processes as part of the complete data science lifecycle (The Team Data Science Process lifecycle, 2022).

## Learning Objectives

After taking this course, the students will be able to:

* Describe fundamental data and database concepts
* Explain and use the database development lifecycle
* Create databases and database objects using popular database management system products
* Solve problems by constructing database queries using Structured Query
* Language (SQL)
* Design databases using data modeling and data normalization techniques
* Develop insights into future data management tool and technique trends Recommend and justify strategies for managing data security, privacy, audit/control, fraud detection, backup and recovery
* Critique the effectiveness of DBMS in computer information systems

## Project Requirements

INSERT PROJECT INSTRUCTIONS

## Project Development

INSERT PROJECT

## Project Reflection

Data collection and subsequent analysis maintained a significant role throughout history and remains vital to the decision-making process today. The science behind understanding that data drives some of the most important advancements through human history, from collecting the positions of stars to fighting the COVID-19 pandemic through information analysis. Companies and governments throughout the world rely on some level of descriptive, predictive, or prescriptive analytics all driven by leveraging data science to derive insights into problems or questions.

The base level of all data starts with a problem or question that could range from a basic “How can we make more money?” to advanced problems like developing self-driving cars. The “science” aspect in data science elicits the idea that undertaking data science problems follows the similar scientific method as many other scientific fields, which all start with asking a question. The initial question or questions of interest drive the acquisition or collection of data for understanding and then provide insight into the modeling and deployment processes as part of the complete data science lifecycle (The Team Data Science Process lifecycle, 2022).

# Course Highlight 4: Applied Machine Learning for Data Analysis

## Course Description

Data collection and subsequent analysis maintained a significant role throughout history and remains vital to the decision-making process today. The science behind understanding that data drives some of the most important advancements through human history, from collecting the positions of stars to fighting the COVID-19 pandemic through information analysis. Companies and governments throughout the world rely on some level of descriptive, predictive, or prescriptive analytics all driven by leveraging data science to derive insights into problems or questions.

The base level of all data starts with a problem or question that could range from a basic “How can we make more money?” to advanced problems like developing self-driving cars. The “science” aspect in data science elicits the idea that undertaking data science problems follows the similar scientific method as many other scientific fields, which all start with asking a question. The initial question or questions of interest drive the acquisition or collection of data for understanding and then provide insight into the modeling and deployment processes as part of the complete data science lifecycle (The Team Data Science Process lifecycle, 2022).

## Learning Objectives

After taking this course, the students will be able to:

* Describe fundamental data and database concepts
* Explain and use the database development lifecycle
* Create databases and database objects using popular database management system products
* Solve problems by constructing database queries using Structured Query
* Language (SQL)
* Design databases using data modeling and data normalization techniques
* Develop insights into future data management tool and technique trends Recommend and justify strategies for managing data security, privacy, audit/control, fraud detection, backup and recovery
* Critique the effectiveness of DBMS in computer information systems

## Project Requirements

INSERT PROJECT INSTRUCTIONS

## Project Development

INSERT PROJECT

## Project Reflection

Data collection and subsequent analysis maintained a significant role throughout history and remains vital to the decision-making process today. The science behind understanding that data drives some of the most important advancements through human history, from collecting the positions of stars to fighting the COVID-19 pandemic through information analysis. Companies and governments throughout the world rely on some level of descriptive, predictive, or prescriptive analytics all driven by leveraging data science to derive insights into problems or questions.

The base level of all data starts with a problem or question that could range from a basic “How can we make more money?” to advanced problems like developing self-driving cars. The “science” aspect in data science elicits the idea that undertaking data science problems follows the similar scientific method as many other scientific fields, which all start with asking a question. The initial question or questions of interest drive the acquisition or collection of data for understanding and then provide insight into the modeling and deployment processes as part of the complete data science lifecycle (The Team Data Science Process lifecycle, 2022).

# Conclusion

The reflections contained within this portfolio report exemplify the significant areas of knowledge gained throughout the Applied Data Science program at Syracuse University. Each of the classes outlined above hold a piece of the axiomatic data science puzzle and represent the foundational blocks for communicating both the statistical processes involved in data analysis and the model development required to reach the predictive and prescriptive elements successful organizations need. Beyond the tools learned in this program, the lessons and strategies developed for achieving solutions derived from data analysis show in the careful development of projects integral for permanent success in the field.

The projects outlined above represent major practice areas in the field of Data Science and highlight the skills obtained in using widely available tools in conjunction with statistical principles. Collection and organization of data using the data warehouse and database principles drive the subsequent analysis and visualization of data through displaying the necessary strategies for handling organizational data. Without these steps, businesses would struggle in achieving functional decision-making and would be unable to take advantage of key insights into business processes.

The theory behind the methods within data science augment the application of data collection, exploration, visualization, analysis, and modeling for enterprise-level operations and are essential to successful decision-making. Each party involved in organizational processes represents an important part of achieving long-term success and communication is integral in serving as the connection point between teams within an organization. FINISH TALKING ABOUT 6 AND 7 OF THE PROGRAM LEARNING GOALS.

# References

*Applied Data Science Master’s Degree.* (2022). Syracuse University. Retrieved February 12, 2022, from [**https://ischool.syr.edu/academics/applied-data-science-masters-degree**](https://ischool.syr.edu/academics/applied-data-science-masters-degree)

Block, Gregory. (2021, June 30). *Applied Machine Learning for Data Analytics.* Syracuse University, IST 707.

Fox, Jonathan. (2021, September 29). *Big Data Analytics.* Syracuse University, IST 718.

Harper, Chad. (2021, January 4). *Database Administration Concepts and Database Management.* Syracuse University, IST 659.

Khan, Humayun. (2022, January 3). *Data Warehouse.* Syracuse University, IST 722.

Larche, Michael. (2021, September 29). *Natural Language Processing.* Syracuse University, IST 664.

Stinnett, John. (2022, January 3). *Project Portfolio Milestone Requirement.* Syracuse University, IST 782.

*The Team Data Science Process lifecycle.* (2022, February 11). Microsoft. Retrieved February 12, 2022, from [**https://docs.microsoft.com/en-us/azure/architecture/data-science-process/lifecycle**](https://docs.microsoft.com/en-us/azure/architecture/data-science-process/lifecycle)