|  |
| --- |
| MS in Applied Data Science Portfolio Submission  William Jones (SUID - 472812314)  wjones01@syr.edu/wsjones4193@gmail.com  Syracuse University |

**Table of Contents**

2 - 3 Resume

4 – 6 Written Evidence

7 Course List & Results

7 References

**William Jones, CPA**

wsjones4193@gmail.com

331 Grove Street

Melrose, MA 02176

404-775-0169

Education

**Syracuse University** Syracuse, NY

**School of Information Studies**

*Master of Science in Applied Data Science*

GPA: **3.909** *January 2020 – Current (Expected June 2022)*

* Relevant coursework: Data Administration Concepts and Database Management, Data Analysis and Decision Making, Business Analytics, Introduction to Data Science, Marketing Analytics, Data Analytics in R, Data Mining, Advanced Information Analytics, Information Visualization, Quantitative Reasoning, Information Policy

**University of Massachusetts Amherst** Amherst, MA

­**Isenberg School of Management**

**­***Bachelor of Business Administration in Accounting* *Graduate, May 2015*

­­Minors: Economics; Education

­Cumulative GPA: **3.47**; Dean’s List 3 Semesters

Skills/Certifications

*Computer*: SQL, Microsoft SQL Server Management Studio, R, RStudio, DAX, Microsoft Power BI, Microsoft Office Applications

*Certifications:* Certified Public Accountant (CPA) registered with the American Institute of CPAs (AICPA) and Massachusetts Society of CPAs (MSCPA)

PROFESSIONAL EXPERIENCE

**BDO USA, LLP** Boston, MA

Manager/Sr. Manager – Audit Data Analytics *January 2021 - Current*

* Developed general (application like products available for use by the 4,000+ BDO audit clients) and custom (bespoke products designed specifically to solve business problems for single or small group of homogenous clients) solutions for audit engagement teams to improve the efficiency and effectiveness of their audits utilizing tools such as PowerBI, Alteryx, SQL, and Python.
* Lead an ever-growing team of 10-20 onshore and offshore data scientist. Responsibilities include project management and staffing as well as upskilling of team members (much of our team has come from an accounting background with minimal data science/analysis experience). Team has produced over 20+ general and 30+ custom solutions and maintains a robust backlog of projects for future development.
* Liaised with client service professionals during data acquisition and development phases of solution development. Often this included significant coaching around databases and storage during the acquisition phase and data modeling concepts during the development phase.

**RSM US, LLP** Boston, MA

­*Assurance Supervisor* *December 2019 – December 2020*

* Audit Data Analytics Center of Excellence IDEAScripting Champion. Created general purpose and custom automation and analytics applications using IDEA and IDEAScripting in an effort to improve the efficiency and effectiveness of RSM’s audits.
* Performed engagement in-charge role on financial statement audit engagements for real estate investment funds. Engagements included commercial real estate clients under US GAAP, IFRS, and OCBOA standards.

**BDO USA, LLP** Various

­*Assurance Associate/Senior Associate* *October 2015 - August 2017, January 2018 – November 2019*

* Audit Data Analytics. Participated in rotation with the National Assurance team building data analytic tools, models and dashboards in Microsoft Power BI to be used by BDO USA and Global audit engagement teams. Rotation included educating senior leadership on the value of data analytics and administering trainings to varying levels of BDO professionals on their use.
  + Performed engagement staff and in-charge role on financial statement audit engagements for a diverse set of clients. Engagements included commercial real estate clients under US GAAP, IFRS, and OCBOA standards.

**Wolf and Company, PC** Boston, MA

­*Senior Auditor* *September 2017 – December 2017*

* Worked on a variety of specialized filings for publicly traded companies (S-1 Filings, S-3 Filings, Comfort Letters, etc.).
* Analyzed and documented the company’s internal control structure in relation to the control environment understanding and for the purposes of audit reliance.

,

Collegiate ACTIVITIES

|  |  |
| --- | --- |
| ­­**University of Massachusetts Athletic Department**  ­ *NCAA Division 1 Baseball Player - Catcher* | ­*Fall 2011 - May 2013* |

* ­­Balanced a 20 hour per week practice and game schedule while keeping a 3.47/3.16 GPA as an Accounting major.
* Participated in the Student Athlete Advisory Committee (SAAC) which organized and help run community service events and fundraisers and voted on different issues that effected student athletes.

**Written Evidence**

*“Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from noisy, structured and unstructured data, and apply knowledge and actionable insights from data across a broad range of application domains.” Dhar, V. (2013).* I love this quote as it clearly articulates the multifaceted nature of data science. In this written evidence section, I will go through each of the 7 “Learning Goals” set out by the MS in Applied Data Science program and demonstrate how my coursework as well as professional experience/resume helps evidence my capabilities as a data scientist.

Learning Goals

1. Describe a broad overview of the major practice areas in data science

Data science and data science projects have many different aspects and stages. The first stage of any data science project is understanding the business problem. Understanding what the end user/decision maker is trying to accomplish is key to designing an efficient and effective solution. This will allow for the data scientist to select the correct tools to complete the job as not all data science projects are the same. Once the problem is understood, it is our responsibility as data scientist to acquire the data needed to solve the business problem. Often this can be messy and/or from multiple sources requiring a data scientist to be able to cleanse the data to make it workable. During the data cleansing process, understanding the business problem and eventual solution is essential as there may be some required data cleansing steps to perform the desired analysis.

Once the data has been cleansed and manipulated as needed, the “fun” part of the data science process begins. The “fun” step is the application of what the end user/decision maker typically thinks about when referring to data science. This is the data visualizations or incorporation of predictive modeling. Again, during this stage, understanding the business problem and eventual solution is essential in determining the correct tools and techniques needed. An example of this is, if an end user is trying to understand a population the visualizations such as histograms are helpful to understand distributions or unsupervised models are helpful to understand clusters within a population whereas if an end user is trying to make a prediction about future events based upon historical events a supervised model is best.

The final stage of any data science project is presentation of the results to the end users/decision makers. This stage cannot go underestimated. Great insights cannot be acted upon if the decision makers cannot interpret them. Clear communication isn’t a goal, it is a must have.

To demonstrate my abilities for this competency, I would like to reference the **Coaches Assignment** from my academic history. The Coaches Assignment from IST 718 helps demonstrate my ability to complete a data science project as described above. In this project there was significant data cleaning, required investigation of the underlying data as well as predictive analysis, and included a report to explain the results to an end user/decision maker in plain English with accompanying technical support.

2. Collect and organize data

Data acquisition is often the most difficult aspect of a data science project. This is also an opportunity for a thoughtful data scientist to provide significant value to the project. Data cleansing has become increasingly easy with the development of no-code and low-code applications. While there are still many instances where use of SQL, Python, or R is necessary to manipulate the data, the expansion of applications like PowerBI and Alteryx has enabled many less technical individuals to perform data science functions.

To demonstrate my abilities for this competency, I would like to reference the **Superheroes Projects** from my academic history. In both IST 687 and IST 718 I utilized the same dataset to complete my final assignments. In each of these classes I utilized a different data cleansing method. In IST 618, I utilized R to clean and manipulate the data, as use of R was a main aspect of the class. In IST 718, I utilized PowerBI to do the same data cleaning and manipulation, as the course allowed for whatever was easiest and most efficient for the students. This is a great example of how data science has developed for into a space where programmers and non-programmers can perform data science tasks.

3. Identify patterns in data via visualization, statistical analysis, and data mining

Once the business problem is understood and the data is acquired and cleansed, that is when the fun begins. This is when statistics, data visualization, and data mining techniques are used to gain a deeper understanding about the data before the modeling phase. At times this can be the whole business problem. In a world where more and more companies have yearly transactions in the 100 million or even billion range, understanding what is happening with your business and being able to dive into the noise is becoming increasingly valuable.

To demonstrate my abilities for this competency, I would like to reference the **NBA Player Stats Project** from my academic history. In IST 652 I created a python script which pulled information from basketball-reference.com and created a dataframe for analysis. I then utilized PowerBI and created a dashboard to visualize the data for deeper understanding. Tools such as PowerBI and Alteryx have made data visualization and interpretation significantly easier to create and more dynamic for the end user/decision maker.

4. Develop alternative strategies based on the data

In certain business problems, the data science tool or method of choice is clearly apparent. In other instances, there are many different choices a data scientist must evaluate to create the best model for the task at hand. Finally, there are many instances that what a data scientist should do with the data will only come after statistical, data visualization, and data mining techniques have been applied. A good data scientist has preconceived notions about what type of techniques would be best for the problem but also is open to making on the fly adjustments as needed.

To demonstrate my abilities for this competency, I would like to reference my professional experience/resume. My current role at BDO is the team lead for the Data Science Solutions team for BDO’s (5th biggest international accounting firm) Assurance and Accounting practice. The team I lead often builds custom solutions for client service teams to improve the way they do their financial statement audits. There are many times when the way to complete the project is apparent from the first meeting but often significant investigation about the business environment is needed before being able to craft an appropriate solution. This is where I believe the Data Science Solutions team provides true value to our client service professionals as we are able to interpret their business problem and design a unique solution using a variety of different possible techniques.

5. Develop a plan of action to implement the business decisions derived from the analyses AND 6. Demonstrate communication skills regarding data and its analysis for managers, IT professionals, programmers, statisticians, and other relevant professionals in their organization

Great insights cannot be acted upon if the end users/decision makers cannot interpret them. Clear communication isn’t a goal, it is a must have.

To demonstrate my abilities for this competency, I would like to reference the **Superheroes Projects** from my academic history. In both IST 618 and IST 718 I utilized the same dataset to complete my final assignments. In each of these projects, I provided a way for the end user/decision maker to interpret a data population they were unfamiliar with and then was able to make selection recommendations based upon their defined needs. What made these projects special was the ability to create a dynamic story that drew the audience in and then provided actionable insights based upon their needs.

7. Synthesize the ethical dimensions of data science practice

In a world of increasing data collection and data science driven decisions, it is essential for data scientist to understand the ethical dilemmas at hand, the impact of the recommendations we are making, and act in manner that is moral and just. While it is important to understand that people will have their own opinions on what is right and wrong, it is key to have discussions around ethical dilemmas in the academic arena to allow for data scientists to have well informed opinions on such matters.

To demonstrate my abilities for this competency, I would like to reference the **Covid-19 Disclosure Paper** from my academic history. During IST 618 we were tasked with writing papers on ethical topics relating to data and information policy related topics. My company had just requested for me to disclose my vaccination status and I wanted to write a paper understanding the implications of this request. The research for the paper and overall themes of the class were truly enlightening.

**Course List & Results**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Semester** | **Class Nbr** | **Class** | **Name** | **Grade** |
| Spring 2022 | 37561 | IST-718 | Big Data Analytics | Active |
| Spring 2022 | 37495 | IST-618 | Information Policy | A |
| Fall 2021 | 22073 | IST-772 | Quant Reasoning Data Science | A |
| Summer 2021 | 71075 | ACC-652 | Accounting Analytics | A |
| Spring 2021 | 37727 | IST-652 | Scripting for Data Analysis | A |
| Spring 2021 | 37838 | IST-719 | Information Visualization | B+ |
| Fall 2020 | 16714 | IST-707 | Data Analytics | A |
| Fall 2020 | 17203 | MAR-653 | Marketing Analytics | A |
| Summer 2020 | 70664 | IST-687 | Introduction to Data Science | A |
| Summer 2020 | 72000 | SCM-651 | Business Analytics | A- |
| Spring 2020 | 37809 | IST-659 | Data Admin Concepts & Db Mgmt | A |
| Spring 2020 | 43881 | MBC-638 | Data Anls & Decisn Making | A |

**References**

Coaches Assignment – IST 718

Superheroes Projects – IST 687 & IST 718 (completed through milestone 2)

NBA Player Stats Project – IST 652

Covid-19 Disclosure Paper – IST 618

All documents can be found at my GitHub here: <https://github.com/wsjones4193/IST_Final_Submission>