

# Sebastian Di Francesco

180 Jackson St NE Atlanta #7322, GA 30312

404-889-1216 | [sebastiandifrancesco@gmail.com](mailto:sebastiandifrancesco@gmail.com)

Linkedin: <https://www.linkedin.com/in/sebastian-di-francesco-b9978873/>

GitHub: <https://github.com/sebastiandifrancesco>

## Summary

Data analyst/scientist with skills in Microsoft excel, statistics, python, databases, web technologies and data visualization, cloud computing, and machine learning. Increasingly innovative ability is leveraged to solve complex problems involving big data that is sometimes structured and sometimes raw. High motivation due to intellectual curiosity allows for proficient analysis with great attention to detail. Strengths include finding truth within visualizations of big data, as well as, knowing how to make certain parts of the visualizations draw more attention than other parts of the visualization according to sense and perception psychology.

## Technical Skills

**Microsoft Excel:** VLOOKUP | Pivot Tables | Conditional Formatting | VBA Scripting

**Statistics:** Modeling | Forecasting

**Python Programming:** Pandas | NumPy | Seaborn | Matplotlib | SciPy | API Interactions | Social Media Mining | Jupyter Notebook | Sklearn | Xgboost | Bs4

**Databases:** Microsoft SQL Server | SSMS | T-SQL | MySQL | MongoDB | ETL

**Web Technologies and Data Visualization:** HTML | CSS | Bootstrap | Dashboarding | JavaScript Charting | D3.js | Geomapping with Leaflet.js

**Cloud Computing:** IBM DB2 on Cloud

## Projects

**Predict-NBA** | <https://github.com/sebastiandifrancesco/Predict-NBA.git>

- Data from the 2013-2014 NBA season was scraped from <https://www.sports-reference.com/> and was used to build different machine learning classification models. After finding the best model and tuning it the model was applied to the 2018-2019 NBA season to evaluate the model further.
- Python | Pandas | Seaborn | Sklearn | Matplotlib | Xgboost | Bs4

**Covid-Prediction** | <https://github.com/sebastiandifrancesco/Covid-Prediction.git>

- Pulled data from Google's Big Query from the covid\_19\_open\_data database. The target feature was new hospitalized patients (on a given day) and ultimately a random forest reg model was constructed.
- Python | Pandas | Sklearn | Matplotlib

**NBA-Cluster** | <https://github.com/sebastiandifrancesco/NBA-Cluster.git>

- Player statistics from the 2018-2019 NBA season was explored and ultimately clustered using a Kmeans model. Five different clusters were generated.
- Python | Pandas | Seaborn | NumPy | Matplotlib | Sklearn

## Education

**Boot Camp Certificate: Georgia Institute of Technology** Atlanta, GA Dec 2020 - May 2021

A data analytics program which taught technical programming skills in Excel, VBA, Python, R, JavaScript, SQL Databases, Tableau, Big Data, and Machine Learning.

**Bachelor of Arts in Cognitive Science: University of Georgia** Atlanta, GA

May 2020

## Professional Experience

**South Carriage Homes** Bishop, GA

May 2019 – Aug 2019

Assistant Project Manager

- Managed subcontractors which ultimately expedited the construction of different custom houses
- Consulted with builder on how to properly manage water drainage systems around build sites and then implemented proper systems to prevent against any possible rain damage