# ZHUOQUAN CHEN

BROOKLYN, NY | 718-300-0078 | zhuoquan1223@email.com | linkedin.com/in/zhuoquanchen
https://zhuoquan-chen.github.io/portfolio/

I am a data scientist with a background in computer science. Skilled in Python and machine learning, I have a passion for providing effective data science solutions and applications. I am a creative problem solver who is also an engaged team player.

## SKILLS

Languages / Environments | Python, SQL, Java, C++.
Packages | Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Plotly,
Modeling | Statistical Modeling, Regression Models, Classification Models, Time Series Analysis, Neural Networks, Image Classification.
Statistics | Statistical Distributions, Bayesian Analysis, p-Values, Hypothesis Testing

## PROJECTS

### STOCKS PREDICTION

Stocks prediction project contains analysis and price prediction. In the analysis part, I analyzed trends and returns of stocks. In the price prediction part, I used the time series model ARIMA and find the best p, d, q to fit the model for predicting the price movements.

- Return calculation • Moving windows • Volatility • Trading strategy: Momentum strategy • Modeling & evaluation

### COVID-19 CASE COUNT ANALYSIS

The primary aim of this analysis is to highlight the key factors that contributed to confirmation of Covid-19 cases in the United States. The goal is to create a model that can be used to predict future Covid-19 cases and cases of a future disease of similar magnitude.

- Analyzed the relationships between features in each dataset.
- Combined those features, which have significant correlation.
- Compared the mask usage status in the United States with New York State.
- Using Choropleth maps to display comprehensive information of each state of the United States and each county of New York.
- Built models for prediction in confirmed cases.

### GAMES CLASSIFICATION

Scraped game posts from Reddit and developed a Natural Language Processing model that identifies which content of posts belongs to board games and which content of posts belongs to card games.

- Applied CountVectorizer to tokenize all content.
- Applied models include Naive Bayes and Gradient Boosting Regression Tree.
- Applied GridSearch to find the hyperparameters for improving models.
- Applied Confusion matrix for evaluating the performance of all models.

## EDUCATION

### GENERAL ASSEMBLY

Data Science Immersive Student, September 2020 - December 2020
12-week full-time immersive educational program strengthening Data Science skills including: Python, SQL, data cleaning, data visualization, regression models, classification models, web-scraping, APIs, NLP, advanced supervised learning, unsupervised learning, time series analysis, and statistics.

### BROOKLYN COLLEGE | NEW YORK, NY

B.S. in Computer Science | Graduated May 2020