

# YUXIN (DIANA) ZHANG

550 Bloomsdorf Dr, State College, PA 16801  
(717) 487-1377 | yuxinzhang9615@gmail.com  
Web: <https://zyx9615.wixsite.com/yuxinzhang>

## OBJECTIVE

To obtain a full-time job or internship in data science / business analysis / statistics programming field.

## EDUCATION

### The Pennsylvania State University, University Park, PA

#### Bachelor of Science in Applied Statistics

Expected Graduation: 05/2019

#### Relevant Courses:

- STAT 380 – Data Science Through Statistical Reasoning and Computation
- DS 310 – Machine Learning • STAT 463 – Applied Time Series • STAT 414 – Probability Theory

## PROJECTS EXPERIENCES

### Summer Research – **Group Leader**

Objective: Develop the Book of Apps for Statistics Teaching

6/2017 - present

- Designing the user interface and using CSS/HTML to style the apps
- Coding in R language to build apps and dashboards
- Maintaining the apps and adapting them to mobile phone
- Presenting in the showcase and exhibition at the Penn State Statistics Department
- Conducting experiments during class sessions, building surveys, collecting data, and analyzing the feedback

### Police Data Challenge – **Group Leader**

Objective: Give suggestions to Baltimore Police on improving public safety

10/2017

- Exploring the 911 call service data, seeking external data and merging datasets
- Building SARIMA time series model using R
- Visualizing the data using Leaflet, ggmap, R Shiny, and ggplot2

### DataFest Competition – **Group Leader**

4/2017

Objective: Analyze booking versus browsing rates on Expedia

- Accessing and analyzing big data in industry with size larger than 2 gigabytes
- Working intensely with teammates for 32 hours in two days

### Class Projects – **Individual / Group Project**

11/2017

- Kaggle – Digits recognition (deep learning and image classification) – accuracy 0.995
- Kaggle – Building a music recommendation system for KKBox's Music – top 23%
- Clickbait Detection using Machine Learning – Using R, Python on feature engineering and Weka on classification to detect clickbait from tweet data
- Monte Carlo – Applying R in random numbers generation and data simulation
- The Study of State College Temperature Over Years – Analyzing time series data in R
- Study cross-sex friendships – Designing survey and conducting statistical analysis

11/2017

9/2017

2/2017

1/2017

4/2016

## SKILLS & CERTIFICATIONS

### Technical Skills:

- Programming:  R (advanced)  R Shiny(expert)  Python(proficient)  CSS/HTML(basic)
- Tools:  TensorFlow  MXNet  Microsoft Office (expert in Excel)  Weka  Power BI  
 SAS(proficient)  Tableau(proficient)  Azure  Photoshop
- Concepts:  Machine Learning  Deep Learning  CNN  Text mining  Regular expression  
 NLP  Classification  Clustering  Naïve Bayes  Regression  Monte Carlo

### Certification:

- Microsoft Professional Program in Data Science (in progress)

### Languages:

- English (fluent)  Chinese (native)

## LEADERSHIP & WORKING EXPERIENCES

### Research Assistant, Penn State Department of Energy and Mineral Engineering

12/2017 - present

### Research Assistant, Penn State Department of Statistics

6/2017 - present

### Online teaching host, English speech training webstream

6/2016 - 7/2016