# Xia Fu

(917) 631-6784  |  xiafu95@gmail.com  |  https://xialalala.github.io/index.html  |  www.linkedin.com/in/xia-charlene-fu

## EDUCATION

**Master of Science (Business Analytics, Data science & Analytics), GPA: 3.947/4.0**          *Aug. 2019 – May 2020*
Rensselaer Polytechnic Institute (RPI)          *Troy, NY*
Coursework taken: Data mining, Machine Learning Applications, Statistics for Managerial Decision Making, Adv Quantitative Methods, Business Analytics Capstone, Adv Artificial Intelligence/Machine Learning for Finance

**Master of Science, Cell and Molecular Biology**          *Aug. 2017 – Aug. 2019*
Fordham University (Fordham University GSAS Centennial Scholarship)          *Bronx, NY*

**Bachelor of Science, Basic Medicine**          *Aug. 2013 – Jul. 2017*
Capital Medical University          *Beijing, China*

## DATA SCIENCE PROJECTS

**Hotel Booking Cancellation Analysis (Python)**          *June 2020*
- Conducted deep EDA analysis to understand each feature of the dataset and did data cleansing and missing value imputation.
- Put forward questions to ask and built models

**Health Home Services Analysis, New York State Department of Health** *(Python, Tableau)*          *Jan. – May 2020*
- Identified abnormal trends through visualizations and provided improvement insights to inform policy on Health Home program
- Designed and developed the Tableau Dashboard to realize real-time monitoring health homes' performance.
- Understood, analyzed and summarized large and diverse sources of data into meaningful and insightful presentations.

**Home Credit Mortgage Default Risk Prediction** *(Python)*          *Jan. – May 2020*
- Exploratory data analysis: data visualizations and summarized analysis to determine the types of insights derived from the datasets.
- Joined together different datasets, conducted data cleansing and built unsupervised machine learning models.
- Applied critical thinking and analytical skills to troubleshoot issues, suggested improvements.

**COVID-19 Data Analysis and Its Impact on Financial Markets** *(R, R markdown Notebook)*          *April 2020*
- Scraped COVID-19 data from JHU databases and financial market returns by quantmod package.
- Used R to visualize the spread of COVID-19 on a world map (Leaflet)
- Built regression model and decision tree model to predict the market trends based on the COVID-19 data

**CFPB consumer complaints analysis** *(Python, NLP)*          *April 2020*
- Searched through the consumer text narratives by NLP to identify the top issues of consumer complaints.
- Provided advice to the federal government and proposed specific pieces of legislation to the US Congress.

**Airbnb New User Bookings Prediction Analysis (Python)**          *Sep. 2019 – Dec. 2020*
- Designed and trained classification models to predict the first destination of new users
- Determined the important factors that drives people's decisions and determined the classification results.

**Gene Knockout Analysis (Python, Galaxy)**          *June 2019*
- Designed and implemented the AURKA Gene knockout experiment.
- Preformed sequence matching and validation; ensured sequence security and integrity measures are met.

## LEADERSHIP EXPERIENCE

**Co-Founder  --  EIPO Consulting Group**          *May 2020 - Present*
- Managing 1000+ members group; Providing free big data and AI-related consulting services to small businesses to help them keep up with current trends.

## OTHER EXPERIENCE

**Market Analyst -- InSitu Composites (NSF I-Corp Program)**          *Oct. 2019 – April 2020*
- Performed valid market research SWOT analysis. Conducted client interviews to stay on top of market needs & pain points in order to help the product team to build the roadmap.

**Product Analyst -- Ehomie New York INC. (ERP system, Product Design)**          *Dec. 2019 – Jan. 2020*
- Collected and analyzed user's behavior data from the ERP system, partnered with product and IT teams to find strategies to improve users' experience, impression and conversion rate.

## TOOL & SKILLS

**Tools:** Python, R, SQL, Tableau, SSIS, SSAS, SSRS, SPSS, Google Analytics

**Skills:** Data Cleansing, Exploratory and Predictive Analysis (Logistic Regression, Random Forest, Boosting Tree, Neural Networks), Clustering, NLP, Statistical Inference (Bayesian Data Analysis), Experiment Design & Testing, Dashboard designing