

## OBJECTIVE

---

Master degree of **Statistics & Data Mining Program**, qualified for career opportunities where **Statistician and SAS Programmer** background will be of value.

## KEY QUALIFICATIONS

---

- 2 years' academic and project experience in **Statistician and SAS Programming**
- 1+ years' project experience in **Biostatistics** sdf
- six months project experience in **Time Series Analysis**
- Well-versed in statistical tools: **SAS (Base, Enterprise Miner, Proc SQL, Graph, Insight, STAT, ODS, Macro Language), R/S-PLUS, SPSS, JMP, Excel, Access, PowerPoint, and AutoCAD**

## EDUCATION

---

<b>Master of Science in Statistics &amp; Data Mining</b> University of Central Florida (UCF), FL	08/2008-05/2010
<b>Master of Science in Material Science &amp; Engineering</b> University of Central Florida, FL	08/2006-05/2008
<b>Bachelor of Science in Mechanical Engineering,</b> Kunming University of Science & Technology, China,	08/1998-07/2002

## AWARD

---

Passed the test of " <b>SAS Global Certification Program</b> "	01/2009
--	---------

## RESEARCH & WORK EXPERIENCE

---

- **Bio-statistician, Biology Department, University of Central Florida,** 01/2009-Present  
Performed statistical analysis concerning "The effect of climate change on the phenology and distribution of nonnative Class I and II EPPC plants in Florida". The statistical methodologies included **Time Series Analysis, Cluster Analysis,** and **Multiple Regressions**. The analysis results will be published at the magazine of "Science".  
Technologies: **SAS/Base, SAS/Macro Language, SAS/Graph, SAS/Insight**
- **Business Project "Predicting Potential Donors in 2010", UCF Foundation Department**  
09/2009-12/2009  
Provided thorough research regarding the donors of UCF, utilizing relevant information since 1999. The goal of this project was to build an efficient model to maximize the fund raising strategy. Duties involved **experimental design, data preparation, model building, and forecasting**. The **misclassification rate for the final model was 0.03%**.  
Technologies: **SAS/Enterprise Miner, SAS/Decision Tree, Bagging, Bootstrapping, and Logistic Regression**.
- **Competition Project "Data Mining Shootout 2009: Best Locations for Planting Energy-Grass",**  
**Sponsored by SAS, CMU-RC, and the Dow Chemical Company** 09/2009-12/2009  
Created models of yield for wheat, corn, soybean, and Energy-Grass by county with soils and weather data (of corresponding years). By calculated and compared the profit contribution per acre of crops for the scoring dataset, the best locations for planting Energy-Grass had been determined.  
Technologies: **SAS/Enterprise Miner, SAS/Base, Multiple Regressions, Neutral Network, SAS/Macro, SAS/GRAPH, and SAS/Insight**

- **Data Mining Project “Forecasting Customers Up Sell”**

08/2009

Built a model with the highest predictability to find the characteristics of customers who can be up sold in the next quarter. Performed comprehensive data mining analysis (230 variables & 40,000 observations) that involved **data preparation, market basket analysis, cluster analysis, interaction detection, and dimension reduction**. The area of ROC curve for the final model was 0.9 which strongly demonstrated that the model is good.

Technologies: **SAS/ Enterprise Miner, SAS/Base, Logistic Regression, Decision Tree, Principle Component Analysis**

- **Statistical Computing Project “Robust Singular Value Decomposition (RSVD)”**

12/2009

Programmed an **R code** to perform the **RSVD**. In this algorithm, the basic idea was trying to search the minimized value using the weighted median as the minimum stable converged estimation. The benefit of the **RSVD** is that data with experimental flaws, outliers and missing values can be examined more easily.

Technologies: **R, Weighted Median, SVD**

- **Project “Analysis of L.A. HEART DISEASE DATA”**

04/2009

This data source was from an epidemiological heart disease study on Los Angeles County employees which was a retrospective biomedical study. Performed **Logistic Regression** model to this project by SAS programming. Interpreted the statistical results to the medical fields.

Technologies: **SAS/Enterprise Miner, SAS/Base, Logistic Regression**

- **SAS Lab Assistant, Data Mining Lab, UCF.**

08/2008 – Present

Responsible for consultant on SAS

Responsible for SAS installation

Responsible for data mining projects using SAS/R programs

Responsible for teaching assistant on statistical methods

- **Research Assistant, Advanced Materials Processing & Analysis Center, UCF**

08/2006 – 05/2008

Designed experiments to synthesis and characterize the polymer-derived porous ceramics.

- **Assistant Engineer, Central Iron and Steel Research Institute, Beijing, CHINA.**

08/ 2002 –03/ 2005

Studied on nano materials’ microcosmic structure and test their physical properties.

## **RELEVANT COURSES**

---

- Advanced Computer Processing of Statistical Data
- Theoretical statistics I &II
- Statistical Computing
- Data Preparation
- Linear Regression
- Logistic Regression
- Time Series Analysis
- Data Mining I &II