Raul Contreras Tempe, AZ 480-492-3030 hrcontre@asu.edu https://hectorraul83.wixsite.com/raulcontreras /raul-contreras

PROFESSIONAL PROFILE

Master in Business Analytics (MSBA) with 4.0 GPA and Certified Tableau Specialist with over seven years of experience in technology, data analysis, database management, and project management seeking to increase data efficiency and driving innovative decision-making solutions applying data science methods and techniques. Skilled in machine learning, statistics, data visualization, programming, and business intelligence.

SKILLS

- Machine Learning
- Statistics
- Experimental Design
- Big Data
- Data Visualization
- Data Mining
- Predictive Analytics
- Agile, Scrum
- Deep Learning
- Data Warehousing
- Marketing Analytics
- Business Intelligence

SOFTWARE

- SPSS, Azure Machine Learning, @RISK, Hadoop, Decision Tools, R, Minitab, Python (Jupyter Notebook), MySQL, Tableau, Google Analytics, and Power BI.

EDUCATION

Arizona State University, Tempe, AZ, USA. May 2020
Master of Science in Business Analytics (MSBA).

Instituto Tecnológico de Ciudad Madero, Ciudad Madero, Tamaulipas, México. Dec 2005
Bachelor of Science in Industrial Engineering.

COURSES & CERTIFICATES

- Tableau Desktop Specialist Certification, Tableau Software — May 2020
- Advanced Google Analytics, Google — May 2020
- PMI Agile Certified Practitioner (PMI-ACP). LinkedIn — Feb 2020
- Advanced Microsoft Power BI. LinkedIn — Nov 2019
- Importing Data in Python. DataCamp — Oct 2019
- PH125.2x: Data Science: R Visualization. Harvard University — Sep 2019
- Data Science for Business Leaders. DataCamp — Sep 2019

EXPERIENCE

EHS Data Consultant, Safe T Professionals (Intel Fab 42 assignment) **Feb 2018 - Now**

- Collect, connect, and analyze different datasets in Power BI enhancing effectiveness of the EHS programs by 90%.
- Develop statistical information and business analysis for executives and management increasing 25% of KPIs participation.
- Develop dashboards in Tableau and Power BI from SQL databases and SharePoint to deploy strategies base on data.
- Predict EHS patterns and analytics by managing injury prevention database to improve profit in the company based on a zero (0) incidents culture.

Safety Engineering Manager, Petersen Dean **Mar 2016 - Feb 2018**

- Used descriptive analytics to create reports and dashboards with Tableau decreasing Incident Rate by 60%.
- Presented incident analysis report, predictive analytics, and data visualization to executive team saving supply cost by 40%.
- Managed a team to solve business problems by collecting and analyzing safety engineering requirements to deliver insights and recommendations dropping safety citations by 80%.

HSE Technical Professional, Halliburton **May 2013 - Jan 2015**

- Generated analytical solutions and led implementation of HSE Management System (ISO 14001-9001, OHSAS 18001) reaching 100% of government and client compliance.
- Created statistics and analytics reports to make decisions for the implementation of HSE plans. Safety issues were solved and 100% of the area incident rate was decreased.
- Presented analytical findings to management and created multiple procedures and 5 campaigns resulted from the analysis.

HS&E Specialist I, Baker Hughes **Apr 2011 - May 2013**

- Identified patterns that helped the organization improving by 30% its costs by recognizing new trend opportunities and increasing the organization's competitive advantage through data mining methods and Excel.
- Implemented internal inspections and audits creating data visualizations that rocket participation from 60% to 91%
- Analyzed data sets to develop international statistics for injury prevention achieving better communication in the organization
- Spearheaded implementation of key projects across 10 countries outperforming budget costs by 40%.

Safety Supervisor, Grupo Tampico **Mar 2006 - Oct 2010**

- Developed strategies and created visualizations to effectively convey the information discovered through data analysis keeping Incident Rate reduced by 50% from 2006 to 2010.
- Managed human resources databases and safety programs using SQL reducing insurance costs by 60% with data analytics.

Raul Contreras Tempe, AZ 480-492-3030 hrcontre@asu.edu https://hectorraul83.wixsite.com/raulcontreras /raul-contreras

PROJECTS

Super Bowl Analysis for Fans and Non-Fans

- Used Python to analyze all 54 Super Bowls. Loaded, cleaned, then explored Super Bowl game, television, and halftime show data. Visualized the distributions of combined points, point differences, viewership, and halftime show performances using histograms. Used line plots to see how ad cost increases lagged viewership increases. Discovered that blowouts do appear to lead to a drop in viewers, and there are better ratings when games have high scores.
- Libraries used: Pandas, Matplotlib, Seaborn, Heatmap,

Flight Delay and Cancellation Prediction

- Recommended changes that will reduce delays and cancellations through descriptive, predictive, and prescriptive data analytics.
- Methods used:
    1. Data preparation, Data Exploratory Analysis was executed through Python. Missing values, data type and shape, and analysis of useful features for our model were the initial focus.
    2. Data visualization, visualization was plotted with the help of Tableau. Total flights, delays, cancellations by month, day of week, airport, causes in minutes were analyzed.
    3. Data correlation, exploring the visualizations from Tableau and creating a Multivariate Analysis obtained by using SPSS. Location, Time, and Correlation of multiple variables were summarized.
    4. Predictive Analytic, Binary Classification Model developed in Azure Machine Learning Studio. Smote, Hyperparameter Tuning and Feature Selection approaches were implemented.

Case Study: Hacker Statistics

- Applied visualization in Matplotlib, manipulating dictionaries and data frames in Pandas, boolean logic, control flow, and loops in Python. Used hacker statistics to calculate chances of winning a bet. Used random number generators, loops, and Matplotlib to gain a competitive edge.

Big Data Applications, Stanford Data Lab

- Based on Hadoop and AWS Platform, used Python to build and interpret three popular Big Data applications: Recommendation Systems, Text Mining, and Social Network Analysis models for data analysis (Data Mining – Stanford Data Lab).

Marketing Analysis, Sun Devil Fitness

- Used SPSS and Excel to clean and analyze the survey data with broad range techniques, including descriptive statistics, comparing means, factor analysis, cluster analysis, OLS regression, logistic regression, and conjoint analysis.
- Presented the findings and suggested changes in an ongoing and campaign for the Marketing and Sales department using Marketing Mix Modeling by creating visualizations on Tableau.

Decision Modeling & Design of Experiments

- Designed a Lego race car using 25 full factorial design of experiments with replicates on 4 selected variables to optimize the setting to obtain maximum distance travelled from a ramp. Financial Analysis and model adequacy checks, including residual analysis.
- Modeling on Excel using Solver, data tables, Stat tools, Palisade suite (Precision tree, @RISK) to maximum production, minimize costs.

Customer Satisfaction, Santander Bank

- Used Azure ML and R to train the best classification models for Santander's customer base, predict potential churners, create visualization and spot key elements to improve retention rate and take proactive steps to improve customer's happiness. 370 features/attributes, 76020 instances with imbalance data, and AUC 0.852 (Data Mining/Machine Learning – Predictive Analytics – Kaggle Competition).

Advance Predictive Modeling, Customer Churn Prediction

- Performed text mining on comments in IBM SPSS Modeler for K means and Kohonen nets clustering. Built stacked models with an AUC 0.89 to identify customers likely to churn and suggested promotions for retention.
- Kaggle-Explored, cleaned and processed the data using python, handled imbalanced data using SMOTE, trained and evaluated models (Decision tree, forest, jungle, SVM, Neural nets), tuning to improve performance. (Target Marketing, Insurance claim, House price prediction, Loan default).
- Generated Book Recommendations based on Amazon product co-purchase data using Network Analysis and graph theory.

Raul Contreras Tempe, AZ 480-492-3030 hrcontre@asu.edu https://hectorraul83.wixsite.com/raulcontreras /raul-contreras

Direct Marketing, Portuguese Bank

- Used Azure ML and R to train the best classification models for Portuguese Bank's customer base, predict potential customers, create visualization, and spot key elements to predict if the client will subscribe (yes/no) a term deposit (variable y). 17 features/attributes, 45211 instances with imbalance data, and AUC 0.924 (Data Mining/Machine Learning – Predictive Analytics – Kaggle Competition).

House Prices Prediction, Ames Housing

- Used Azure ML to impute the missing data and do the feature engineering on selected columns.
- Trained the dataset through different regression algorithms like random forest, boosted decision tree and build stacked ensemble models to predict the price of houses in test data (Data Mining/Machine Learning – Predictive Analytics – Kaggle Competition).

Insurance, Fraud Detection, Auto Insurance Company

- Used Azure ML and python to train the best classification models for Auto Insurance Company's customer base, predict if a claim was fraudulent or not. 31 features/attributes, 15900 instances with imbalance data, and AUC 0.921 (Data Mining/Machine Learning – Predictive Analytics – Kaggle Competition).

Porto Seguro's Safe Driver Prediction on Kaggle

- Cleaned and coded the training and test data with over 500,000 rows and 60 features on Excel and Azure. Used SMOTE to remove the imbalance present in the data.
- Used various machine learning algorithms (Decision Tree, Decision Forest, SVM, Neural Networks) and feature selections to select the most important features and ran the model to get better predictions of whether a driver would initiate an insurance claim in the next year or not.

Tableau Visual Analytics Project

- Cleaned and coded multiple data sets using Excel and Python containing sales information from retail as well as e-commerce websites.
- Created charts and interactive dashboards using Tableau to display the results of the analysis with sales trends (in terms of seasonality by type of product) to help develop a better marketing strategy for the company to boost sales.

MySQL Project - Relational Model of apps and app reviews

- Navigated the complex data warehouse using MySQL consisting of many tables to acquire the required data about more than 1,00,000 apps in 6 different classes (such as paid, free) and filtered the top grossing.
- Used nested queries and joins to connect and extract data from multiple tables and to get the top 300 apps of each class in more than 12 different categories (genres) over the course of the year to understand the trends.