Edward Pham

Data Analyst

Seattle, WA | me@edpham.net | 503-706-3519 | www.edpham.net | LinkedIn | Tableau Public

Summary

A data analyst and NLP developer, tackling large data sets and solving problems related to data and language with the mission of improving the healthcare field. An analyst on a podcast covering Major League Soccer, informing listeners and readers to understand the game of soccer with a data-driven perspective.

Technology & Skills

Technologies: Python, SQL (T-SQL and PostgreSQL), Tableau, PowerBI, Javascript, Java, Excel, PowerPoint, Word, Google Sheets
Skills: Data Cleaning, Data Analysis, Data Visualization, Data Transformation

Work Experience

Data Analytics Immersive Fellow | General Assembly | Remote | Oct 2023 – February 2024

- Completed 12-week full-time immersive educational program of hands-on professional training strengthening Data Analytics skills including Python, SQL, data cleaning, data visualization (Tableau and Microsoft Power BI), regression models, classification models, web-scraping, APIs, and statistics.
- Completed multiple projects to solidify learning, which will include a capstone research project.

Research Associate/Analyst | Department of Veteran Affairs | Seattle, WA | May 2014 – January 2018

- Used machine learning algorithms such as Naïve Bayes, k-Nearest Neighbors, and SVM to develop a multi-category classification tool that read unstructured CT report texts from radiologists to determine whether the patient was at risk of lung cancer.
- With the testing set, 29.5% of the scans identified as needing immediate attention with 67% sensitivity and 85% specificity.
- Bridged the gap between programmers and investigators to help understand and communicate issues and concerns about the machine learning study.
- Addressed primary investigators' questions and potential study and grant ideas by utilizing and manipulating data from the VA's SQL databases using T-SQL. Cleaned up the T-SQL data for investigators to simplify their workload for their studies and grants.

Research Intern | GroupHealth Research Institute | Seattle, WA, Jul *2012 – Dec 2012*

- Built a rule-based machine learning classification tool to identify at-risk patients for colon cancer by reading structured electronic medical records using GATE, as well as a post-processing tool with Python. Resulting in 88.5% recall and 98% precision.

Writer/Analyst | Radio Cascadia | Seattle, WA, Sep 2014 – Dec 2018

- Directed data-driven analysis and discussion with colleagues to discuss games in an objective manner to podcast listeners and Twitter followers.
- Used data provided by OPTA and MLS and other analysts to analyze and provide the big picture and details of our post-game analysis during podcast recordings.
- Covered numerous MLS matches on-site including 2015 and 2018 MLS Cup with focus on Portland Timbers.
- Recognized by Major League Soccer, Vancouver Whitecaps, Seattle Sounders, and Portland Timbers.

Projects

NBA Salary Cap/Player Metrics Project | Python | Pandas | BeautifulSoup | Tableau | Data Analysis | Data Visualization

- Utilized Python libraries including Pandas and BeautifulSoup to extract traditional and advanced player metric data from the NBA Stats API. Scraped player salary, salary cap, and luxury tax information from multiple websites to complement the player metrics dataset.

- Conducted in-depth analysis to examine the correlation between player salaries and various performance metrics.
- Developed Tableau dashboards to visually represent the relationship between player metrics and salaries, providing insights into player performance and market value.
- Designed an interactive dashboard to present multiple metrics, facilitating a comprehensive evaluation of player strengths and weaknesses in an accessible format.

Renewable Energy Investment Project | Python | SQL | Data Analysis | Data Visualization | Leadership

- Led a team of three in evaluating the feasibility of a hydroelectric company investing in wind energy within the market.
- Applied Python, Pandas, and SQL to clean and process data effectively. Employed Tableau and Seaborn to craft insightful charts and maps, offering a comprehensive view of potential investment regions. Identified key manufacturers and operators, providing valuable recommendations for strategic decisions in the wind energy market.

Baywheels Project | SQL | Tableau | Database Creation & Management | Data Analysis | Data Visualization

- Examined how the COVID pandemic affected the Baywheels program in the Bay Area.
- Created and managed my own PostgreSQL database and uploaded Baywheels rider data from 2017 until present onto it. Standardized the datasets and created queries to analyze station usage, rider data, membership vs. casual usage, and usage from each city pre-quarantine and post-quarantine.
- Created dashboards to tell a story of the program before and after the pandemic, such as maps of the station usage before and after in San Francisco, San Jose, and Oakland/Berkeley, how ridership increased and decreased after the stay-at home order, and the monthly ridership for the entire program and the rolling average.

AirBnB Project | Excel | PowerBI | Data Analysis | Data Visualization

- Provided a recommendation for potential winter property locations for three families to invest in, using the property's rental income to offset future vacations and the mortgage.
- Analyzed Redfin's Denver historical housing data and Airbnb's historical rental data through Excel to break down potential neighborhoods and rental pricing based on location and size of the property.
- Utilized PowerBI to calculate different KPIs and provide various visualizations to show costs based on neighborhoods, average rental price per night vs. average occupancy rate, and return on investment based on property size.

Education

General Assembly, Data Analytics Immersive Program Fellow, Expected: Feb 2023

Master of Science, University of Washington, Computational Linguistics, Sep 2009 – Dec 2012, *Deep and shallow machine learning techniques, as well as rule and statistical based techniques*.