

## EMPLOYMENT

---



**Daymon Interactions** | San Diego, California  
**Data Scientist**

Nov 2016 - present

- Development of Tableau dashboards, visualizations, and reports of client-specific KPIs.
- Forecasting revenue, predicting employee retention, and data exploration using Python.



**University of Washington eScience Institute** | Seattle, Washington  
**Data Science for Social Good Fellow**

Jun 2017 – Aug 2017

- Identify cruising (vehicles looking for parking and for-hire vehicles) using Seattle's traffic sensor data, develop a pipeline to store, process, and query the data, and determine the temporal and spatial differences in a Seattle cruising heatmap.



**ProQinase GmbH** | Freiburg, Germany  
**Bioinformatician**

May 2016 – Sep 2016

- Data mining of public domain cell line data bases using R (PharmacoGx) and development of a Cancer Cell Line Explorer in Tableau for assessing the RNA expression, mutation, and drug perturbation data of 49,273 genes and 3,768 cells.
- Establishing an interactive Tableau dashboard for customer evaluation of biochemical high-throughput data.

## CERTIFICATIONS

---

**Tableau Server 10 Qualified Associate** | Tableau Software  
**Data Mining** | University of California, San Diego | Extension  
**Python** | University of Michigan | Coursera

Jun 2017  
Jul 2016  
Dec 2015

## COMPETITIONS

---

**Basic Needs** | [github.com/OrysaStus/basicneeds](https://github.com/OrysaStus/basicneeds)

- Developed an SMS-based app for the homeless population which provides the real-time information about nearest locations, routing, weather, and incident reporting. Amazon Lex and Lamba, Twilio, HERE technologies, and Python.
- AngelHack | Seattle, WA | Application Development | **Best Use of HERE Technologies** Jul 2017

**Breaking News** | [newsbreakers.herokuapp.com](https://newsbreakers.herokuapp.com)

- Created a website which takes in a news article user input URL and predicts the probability of fake news based on article text, consistency between the title and article text, and article subjectivity and polarity. Beautiful Soup, Flask, newspaper, SciKit-Learn (Naive Bayes text classification), and textblob.
- Global Artificial Intelligence Hackathon | Seattle, WA | Machine Learning | **Seattle 1<sup>st</sup> Place** Jun 2017

**Data Bots Unleashed** | [wastebotsunleashed.github.io](https://wastebotsunleashed.github.io)

- Extracted a training set of 40 littered and a testing set of ~50,000 random SD Google Street View images. The images were processed, trained using shape and pixel recognition models, classified as littered/not littered, and the littered SD zip codes were visualized. Python (Google Street View API, Pandas, OpenCV, PIL), R (ggmap), Git, HTML, and Tableau.
- San Diego Women's Hackathon | San Diego, CA | Environment | **1<sup>st</sup> Place** Apr 2017

**San Diego Hearts** | [sandiegohearts.github.io](https://sandiegohearts.github.io)

- Designed an interactive website that examines the factors contributing to varying heart disease rates across San Diego. Locations related to heart disease, consumer expenditure, county initiative, and hospital mortality data were used along with Python (YelpAPI, Pandas, TabPy), R (ggmap), Git, HTML, and Tableau.
- Big Data Hackathon | San Diego, CA | Public Health | **2<sup>nd</sup> Place** Feb 2017

## SKILL SET

---

Proficient: Python (Pandas, NumPy, SciPy, SciKit-Learn, StatsModels, Matplotlib, Seaborn, Vincent, Beautiful Soup, PyQt4, Scrapy, NLTK, TabPy), relational databases (PostgreSQL, MS SQL Server, SQLite), Tableau, Git, Spark, TensorFlow  
Other: UNIX Shell Programming, WEKA, R (Bioconductor), D3, SAS

## COLLABORATIONS

---



**Angel Evan Consulting** | Seattle, Washington  
**Data Scientist**

May 2017 - present

- Tableau Server, Tableau Desktop, ETL, modeling, Python, R, and MS SQL for case-to-case data science projects.
- Clients: Skechers, True Religion



**Qualcomm Institute – UCSD** | La Jolla, California  
**Data Scientist**

Jan 2017 - present

- Web scraping 1 million health fundraising campaigns from GoFundMe.com using Scrapy.
- Querying and visualizing data to quantify funding incentives and ethical concerns using Tableau, Seaborn, and Matplotlib, text mining and text classification using multinomial Naive Bayes and NLTK, and predicting probability of funding success.



**Engineering World Health** | La Jolla, CA  
**Synthetic Biology Researcher**

Jun 2015 – Jun 2016

- Aligned 180 sequences from 9 strains of HIV1 and identified 5 highly conserved regions of HIV1 for target identification using Biopython and designed synthetic genetic circuits for sensing HIV RNA.

## FELLOWSHIPS | AWARDS | SCHOLARSHIPS

---

**Data Science for Social Good Graduate Fellow** | Data Science | Microsoft | University of Washington

Apr 2017

**RISE Professional Fellow** | Bioinformatics | DAAD (German Academic Exchange Service)

Mar 2016

**Gordon Engineering Leadership Scholar** | Engineering Leadership | University of California, San Diego

Jun 2015

## TECHNICAL WORK SAMPLE

---

**Seattle Traffic Cruising Analysis** | [github.com/uwescience/TrafficCruising-DSSG2017](https://github.com/uwescience/TrafficCruising-DSSG2017)

- Using traffic sensor data to detect vehicle cruising from individuals looking for parking or for-hire vehicles in downtown Seattle. Some of the tools used include SciKit-Learn, Pandas, NetworkX, RethinkDB, and Leaflet.js.

**GlobalGiving Campaign Analysis** | [public.tableau.com/profile/orysya.stus#!/vizhome/GlobalGiving/GlobalGivingSuccess](https://public.tableau.com/profile/orysya.stus#!/vizhome/GlobalGiving/GlobalGivingSuccess)

- Identifying the campaign success indicators and analyzing the description text vocabulary, sentiment, and readability for successful campaigns for the 2017 Seattle DataDive. Some of the tools used include Pandas, NLTK, textblob, and Tableau.

**Retail Logic for Daymon Worldwide Clients**

- Accessing reported data stored in Redshift, capturing KPIs for retail business success, designing cohesive and interactive dashboards using Tableau, and providing insights for clients.

**The Evolution of Poverty & Income Distribution in the U.S.** | [github.com/OrysyaStus/UCSD\\_Data\\_Science\\_and\\_Engineering](https://github.com/OrysyaStus/UCSD_Data_Science_and_Engineering)

- Developing meaningful data visualizations to depict the evolution of poverty in the U.S. using Matplotlib and Vincent. Employing statistics to gather insights from the data using StatsModels.

**SQL and SAS Reporting** | [github.com/OrysyaStus/UCSD\\_Data\\_Mining\\_Certificate](https://github.com/OrysyaStus/UCSD_Data_Mining_Certificate)

- Establishing a schema and database design for dirty data and creating SQL script for satisfying the employers requests in an efficient manner. Utilizing proc format, means, sort, univariate, and freq along with formats, data set merging, and variable calculations to create a visually appearing SAS report. Data provided by industry collaborates used.

## EDUCATION

---

**University of California, San Diego** | La Jolla, CA

M.S. Data Science and Engineering (online)

Jun 2018

B.S. Biomedical Engineering: Biotechnology

Jun 2016

## RELEVANT COURSEWORK

---

- Python for Data Science Analysis
- Database Management Systems using PostgreSQL
- Probability and Statistics for Data Science
- Machine Learning for Data Science
- Data Analysis Using Spark
- Data Integration and ETL
- Beyond Relational Data Models