DATA SCIENCE CASE STUDY

First, RELAX! This exercise is designed to, first and foremost, gauge your problem-solving skills and ability to work through a project end-to-end – from raw data to business conclusions. And it will also give you the time and space to illustrate your skills.

We will use this case study to determine whether to advance you to the next round of interviews.

So have fun, be creative, and show us what you can do!

INSTRUCTIONS

- Complete the following coding exercise within one week of receiving
- Feel free to use any resources at your disposal (except previously built models), but be prepared to talk through your approach interactively
- Be prepared to discuss:
    - How you framed/scoped the problem
    - Your approach to solving the problem
    - Exploratory data analysis
    - Challenges you faced
    - Things that didn't work
    - How you tested your results
    - The results themselves
    - Business conclusions

DATA

- Dataset comprised of individual times series, one for each imaginary product
- Each time series includes daily units sold, both organic and those resulting from PPC marketing
- Also included is the daily PPC marketing spend, in dollars, for each specific product
- Note that some of the time series have several years of data, others have only a few months

TASK

- Develop time series models to forecast both Organic and PPC sales for each time series
- Build naïve model and models that iteratively improve on the naïve model
- Score models on training, validation, and test data and compare results
- Devise and implement model to improve prediction of products with limited historical data
- Assess impact of PPC Cost on PPC unit sales (i.e. where does PPC have most impact, how to quantify)
- Create short powerpoint or report to describe the problem, your approach, and conclusions
- Prepare to discuss during your next interview

GOOD LUCK!!