**5 Practical Data Science Projects That Will Help You Solve Real Business Problems for 2022**

A curated list of data science projects that mimic real-life problems

“Say goodbye to useless side-projects.”

It’s been almost two years since I started writing articles — that’s equated to just over 175 articles! One fault in some of my previous articles is that I suggested data science projects that were interesting, **but not practical.**

One of the easiest ways to get a job as a data scientist is to show that you’ve already completed similar projects and work as the job posting itself. Therefore, I wanted to share with you some **practical** data science projects that I’ve personally done throughout my career that will beef up your experience and your portfolio.

With that said, let’s dive into it:

**1. Customer Propensity Modelling**

**What?**

A propensity model is a model that predicts the likelihood that someone will do something. To give a few examples:

* The likelihood that **website visitors** will **register an account**
* The likelihood that a **registered user**will **pay and subscribe**
* The likelihood that a **user** will **refer another user**

As well, propensity modeling does not only entail “who” and “what” — it also entails “when” (when should you target the users you’ve identified) and “how” (how should you deliver your message to your targeted users?).

**Why?**

Propensity modeling allows you to allocate your resources more wisely, resulting in greater efficiencies, while achieving better results. To give an example, think of this: instead of sending an email advertisement where there’s a 0%-100% chance of a user clicking it, with propensity modeling, you can target users with a 50%+ chance of clicking it. Fewer emails, more conversions!

**How to:**

Below are two code walkthroughs that demonstrate how to build basic propensity models:

**[Customer propensity to purchase](https://www.kaggle.com/benpowis/customer-propensity-to-purchase" \t "_blank)**

[Explore and run machine learning code with Kaggle Notebooks | Using data from Customer propensity to purchase dataset](https://www.kaggle.com/benpowis/customer-propensity-to-purchase" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/benpowis/customer-propensity-to-purchase" \t "_blank)

**[Marketing Analytics, Classification, and EDA](https://www.kaggle.com/jalenguzman/marketing-analytics-classification-and-eda" \l "Classification-Algorithms" \t "_blank)**

[Explore and run machine learning code with Kaggle Notebooks | Using data from Marketing Campaign](https://www.kaggle.com/jalenguzman/marketing-analytics-classification-and-eda" \l "Classification-Algorithms" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/jalenguzman/marketing-analytics-classification-and-eda" \l "Classification-Algorithms" \t "_blank)

Here are two datasets that you can use to build a propensity model. Take note of the type of features that are offered in each dataset:

**[Customer propensity to purchase dataset](https://www.kaggle.com/benpowis/customer-propensity-to-purchase-data" \t "_blank)**

[A data set logging shoppers interactions on an online store](https://www.kaggle.com/benpowis/customer-propensity-to-purchase-data" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/benpowis/customer-propensity-to-purchase-data" \t "_blank)

**[Marketing Campaign](https://www.kaggle.com/rodsaldanha/arketing-campaign?select=marketing_campaign.csv" \t "_blank)**

[Boost the profit of a marketing campaign](https://www.kaggle.com/rodsaldanha/arketing-campaign?select=marketing_campaign.csv" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/rodsaldanha/arketing-campaign?select=marketing_campaign.csv" \t "_blank)

**Be sure to**[**subscribe**](https://terenceshin.medium.com/membership)**to never miss another article on data science guides, tricks and tips, life lessons, and more!**

**2. Metric Forecasting**

**What?**

Metric forecasting is self-explanatory — it refers to forecasting a given metric, like revenue or the total number of users, in the short-term future.

Specifically, forecasting involves techniques that use historical data as inputs to generate a predicted output. Even if the output itself is not entirely accurate, forecasting can be used to gauge the general trend of where a particular metric is going.

**Why?**

Forecasting is basically like looking into the future. By predicting (with some level of confidence) what will happen in the future, you can make more informed decisions more proactively. The result of this is that you’ll have more time to make decisions and ultimately reduce the likelihood of failure.

**How to:**

The first resource provides a summary of several time-series models:

**[An overview of time series forecasting models](https://towardsdatascience.com/an-overview-of-time-series-forecasting-models-a2fa7a358fcb" \t "_blank)**

[We describe 10 forecasting models and we apply them to predict the evolution of an industrial production index](https://towardsdatascience.com/an-overview-of-time-series-forecasting-models-a2fa7a358fcb" \t "_blank)

[towardsdatascience.com](https://towardsdatascience.com/an-overview-of-time-series-forecasting-models-a2fa7a358fcb" \t "_blank)

The second resource provides a step-by-step walkthrough in creating a time-series model using Prophet, a Python library built by Facebook specifically for time-series modeling:

**[Time Series Analysis and Forecasts with Prophet](https://www.kaggle.com/elenapetrova/time-series-analysis-and-forecasts-with-prophet" \t "_blank)**

[Explore and run machine learning code with Kaggle Notebooks | Using data from Rossmann Store Sales](https://www.kaggle.com/elenapetrova/time-series-analysis-and-forecasts-with-prophet" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/elenapetrova/time-series-analysis-and-forecasts-with-prophet" \t "_blank)

**3. Recommendation Systems**

**What?**

Recommendation systems are algorithms with an objective to suggest the most relevant information to users, whether that be similar products on Amazon, similar TV shows on Netflix, or similar songs on Spotify.

There are two main types of recommendation systems: collaborative filtering and content-based filtering.

* **Content-based**recommendation systems recommend particular items based on previously chosen items’ features. For example, if I watched a lot of action movies previously, it would rank other action movies higher.
* **Collaborative filtering** on the other hand filters items that a user might like based on the reactions of similar users. For example, if I liked Song A and someone else liked Song A **and**Song C, then I would be recommended Song C.

**Why?**

Recommendation systems are one of the most widely used and most practical data science applications. Not only that, but it also has one of the highest ROIs when it comes to data products. [It’s estimated that Amazon increased its sales by 29% in 2019 specifically due to its recommendation system.](https://rejoiner.com/resources/amazon-recommendations-secret-selling-online/" \l ":~:text=%E2%80%9CJudging%20by%20Amazon's%20success,%20the,the%20same%20time%20last%20year." \t "_blank) As well, [Netflix claimed that its recommendation system was worth a staggering $1 BILLION in 2016](https://www.businessinsider.com/netflix-recommendation-engine-worth-1-billion-per-year-2016-6" \t "_blank)!

But what makes it so profitable? As I alluded to earlier, it’s about one thing: **relevancy**. By providing users with more **relevant** products, shows, or songs, you’re ultimately increasing their likelihood to purchase more and/or stay engaged longer.

**How to:**

**Resources and Datasets**

**[Introduction To Recommender Systems- 1: Content-Based Filtering And Collaborative Filtering](https://towardsdatascience.com/introduction-to-recommender-systems-1-971bd274f421" \t "_blank)**

[How services like Netflix, Amazon, and Youtube recommend items to the users?](https://towardsdatascience.com/introduction-to-recommender-systems-1-971bd274f421" \t "_blank)

[towardsdatascience.com](https://towardsdatascience.com/introduction-to-recommender-systems-1-971bd274f421" \t "_blank)

**[Netflix Movies and TV Shows](https://www.kaggle.com/shivamb/netflix-shows" \t "_blank)**

[Movies and TV Shows listings on Netflix](https://www.kaggle.com/shivamb/netflix-shows" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/shivamb/netflix-shows" \t "_blank)

**[Restaurant Recommendation Challenge](https://www.kaggle.com/mrmorj/restaurant-recommendation-challenge?select=orders.csv" \t "_blank)**

[Recommendation Challenge data](https://www.kaggle.com/mrmorj/restaurant-recommendation-challenge?select=orders.csv" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/mrmorj/restaurant-recommendation-challenge?select=orders.csv" \t "_blank)

**[Spotify Recommendation](https://www.kaggle.com/bricevergnou/spotify-recommendation?select=yes.py" \t "_blank)**

[200 songs and their statistics](https://www.kaggle.com/bricevergnou/spotify-recommendation?select=yes.py" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/bricevergnou/spotify-recommendation?select=yes.py" \t "_blank)

**4. Deep Dive Analyses**

**What?**

A Deep dive analysis is simply an in-depth analysis of a particular problem or topic. They can be **explorative** in nature, to discover new information and insights, or **investigative**, to understand the cause of a problem.

It’s not a widely talked about skill, partially because it comes with experience, but that doesn’t mean you can’t improve it! Like anything else, it’s just a matter of practice.

**Why?**

Deep dives are essential for any data-related professional. Being able to figure out why something doesn’t work, or being able to find the silver bullet, is what differentiates great from good.

**Resources and datasets**

Below are several deep dive tasks that you can try on your own:

**[Healthcare cost](https://www.kaggle.com/ravichaubey1506/healthcare-cost" \t "_blank)**

[Analyze the Healthcare cost and Utilization in Wisconsin hospitals](https://www.kaggle.com/ravichaubey1506/healthcare-cost" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/ravichaubey1506/healthcare-cost" \t "_blank)

**[IBM HR Analytics Employee Attrition & Performance](https://www.kaggle.com/pavansubhasht/ibm-hr-analytics-attrition-dataset" \t "_blank)**

[Predict attrition of your valuable employees](https://www.kaggle.com/pavansubhasht/ibm-hr-analytics-attrition-dataset" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/pavansubhasht/ibm-hr-analytics-attrition-dataset" \t "_blank)

**[Marketing Analytics](https://www.kaggle.com/jackdaoud/marketing-data/tasks?taskId=2986" \t "_blank)**

[Practice Exploratory and Statistical Analysis with Marketing Data](https://www.kaggle.com/jackdaoud/marketing-data/tasks?taskId=2986" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/jackdaoud/marketing-data/tasks?taskId=2986" \t "_blank)

**5. Customer Segmentation**

**What?**

Customer segmentation is the practice of dividing a customer base into several segments.

The most common type of segmentation is by demographic, but there are many other types of segmentation including geographic, psychographic, needs-based, and value-based.

**Why?**

Segmentation is extremely valuable to a business for several reasons:

* It allows you to conducted more targeted marketing and deliver more personalized messaging to each segment. Young teenagers value much different things than parents of several kids.
* It allows you to prioritize particular segments when resources are limited, particularly those that are more profitable.
* Segmentation also serves as a basis for other applications like upselling and cross-selling.

**How to:**

**[Customer Segmentation](https://www.kaggle.com/fabiendaniel/customer-segmentation" \t "_blank)**

[Explore and run machine learning code with Kaggle Notebooks | Using data from E-Commerce Data](https://www.kaggle.com/fabiendaniel/customer-segmentation" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/fabiendaniel/customer-segmentation" \t "_blank)

**[Customer Segmentation (K-Means) | Analysis](https://www.kaggle.com/kushal1996/customer-segmentation-k-means-analysis" \t "_blank)**

[Explore and run machine learning code with Kaggle Notebooks | Using data from Mall Customer Segmentation Data](https://www.kaggle.com/kushal1996/customer-segmentation-k-means-analysis" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/kushal1996/customer-segmentation-k-means-analysis" \t "_blank)

**Datasets**

**[Mall Customer Segmentation Data](https://www.kaggle.com/vjchoudhary7/customer-segmentation-tutorial-in-python" \t "_blank)**

[Market Basket Analysis](https://www.kaggle.com/vjchoudhary7/customer-segmentation-tutorial-in-python" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/vjchoudhary7/customer-segmentation-tutorial-in-python" \t "_blank)

**[Customer Segmentation Classification](https://www.kaggle.com/kaushiksuresh147/customer-segmentation" \t "_blank)**

[Classify the customers into four segments](https://www.kaggle.com/kaushiksuresh147/customer-segmentation" \t "_blank)

[www.kaggle.com](https://www.kaggle.com/kaushiksuresh147/customer-segmentation" \t "_blank)

**Thanks for Reading!**

I hope you find this useful in your data science journey! Mastering these skills does not only show that you’re technically sound, but it also shows that you know how to conduct data science projects **that add business value**. Say that to any hiring manager and I guarantee they’ll be impressed ;).

As always, I wish you the best in your learning endeavors!