I was born in India and came to Canada in 2016 to pursue my higher education. Canada is a beautiful country with a diverse range of natural landscapes, and there are plenty of ways to explore and enjoy them. During the winter months, I’ve hit the slopes in Whistler, Vernon, and most recently in Banff. I’ve also ice-skated on the frozen lake louis. In the summer, I try to do bit of hiking (Grouse grind is a workout and a half!) and enjoy water sports. No matter what time of year it is, Canada's natural beauty is always worth exploring. <br><br>

I'm a big fan of staying active and trying new things, so I really enjoy weightlifting and other forms of exercise. I also love spending time with friends and family, whether it's playing board games or just hanging out. Reading is one of my favourite pastimes, and I always have a book or two on the go (currently diving into the deep world of Brandon Sanderson). I'm also an adventurous person and love to travel and explore new places. Whether it's a short road trip or a longer international trip, I'm always up for an adventure. My travel goal is to visit the modern seven wonders of the world (current progress: 2/7).

I am also into playing literally all kinds of sports. I have played national tournaments in tennis and basketball back in India, so feel free to invite me to any game session. I love a challenge and having a healthy competition!!

I am a self-motivated, enthusiastic team worker with a keen interest in Machine Learning and Data Science. One of my greatest strengths is my ability to work well in a team environment. I understand the importance of collaborating with others and strive to contribute my skills and knowledge in order to achieve our shared goals. I am passionate about solving complex business problems through data storytelling and helping make jobs better and more efficient by understanding and supporting the organizational direction. I enjoy being strategic with data and I excel at realizing what needs to be acquired to increase the value of the existing data.<br> <br>

In addition to my passion for data science, I also have a strong interest in sports analysis and fintech. I enjoy using data and analytics to understand and predict trends in these areas, and am always seeking out opportunities to learn and grow in these fields.<br> <br>

I enjoy working directly with clients and establishing meaningful and transparent relationships in order to foster mutual growth. I love learning about what makes an individual business succeed and its business model in order to see where I can help. <br><br>

During my time at SFU, I was part of a team that created a podcast about mental health specifically targeting university students. We each shared our struggles with mental health while attending university and shared our coping strategies to overcome our difficulties. We took a evidence based approach to give guidelines and advice on how to enjoy time at university. I also interviewed a psychology professor that advocates art therapy for depression and the benefits for having creative outlets in our lives.

Customer churn is always a top priority of any subscription based company. Ideally churn should 0.00% but that's only if we lived in an ideal world where unicorns existed. However, in reality, churn should be as low as possible because it costs more to acquire new customers than it is to retain existing customers. This can be especially problematic for businesses that rely on a steady stream of new customers to keep their revenue stream flowing. So it becomes really important for companies to maintain a loyal customer base and simultaneously determine the factors that affect the experience of the customer resulting in them churning. Having that information can help a company make data driven decisions to keep churn rate minimal. High rates of customer churn can be a sign that there are problems with the company's products or services, which can be addressed through targeted improvements.

This project tries to determine the factors that affect churn. It involves predicting whether a customer is likely to switch to a different service provider. Classification algorithms, such as random forest, K-nearest neighbour, logistic regression, and support vector machines can be used to predict telecom churn.

The term deepfakes originated in 2017 when a Reddit user named “deepfakes” shared the first ever of this media online. These videos involved celebrities’ faces swapped onto the bodies of actresses in pornographic videos. Since then academic research and interest in deepfakes (a portmanteau of "deep learning" and "fake") has grown quickly. Deepfakes leverage powerful machine learning to manipulate a source image and generate a video using a driving media to generate visual content with a high potential to deceive. Deepfakes have garnered widespread attention for their uses in celebrity pornographic videos, fake news, hoaxes, and financial fraud. This has elicited responses from both industry and government to detect and limit their use. <a class="link" href="https://colab.research.google.com/github/JaumeClave/deepfakes\_first\_order\_model/blob/master/first\_order\_model\_deepfakes.ipynb" target="\_blank">This project</a> uses Aliaksandr Siarohin First Order Model Method to generate and create deepfakes of various politicians and famous sports figures. The project explains deepfakes and what they are. How they can be created and the technology behind them including generative adversarial networks, neural networks and autoencoders. The paper ends with a section about the ethics of these media and the importance of recognizing that they exist so that the public is less likely to be deceived.

Spotify is a digital music, podcast, and video streaming service that was launched in 2006. It allows users to browse and stream a wide variety of music, podcasts, and other audio content, as well as create and share playlists. Spotify is available in more than 92 countries and has over 345 million active users, making it one of the most popular streaming services in the world. These days, Spotify and music have become synonymous. I use Spotify daily and I feel their recommendation algorithm is a cut above the other music streaming service given their price point of just $10 a month (having tried Apple music, amazon prime unlimited). I always wondered how did they make such accurate recommendations given my streaming behaviour. I wanted to tackle this from a data science perspective, so I set out to create some playlists for myself. This project uses Spotipy, Spotify’s API, to access the top 1000 songs from 2017 to 2022. It takes a detailed look into how a tracks audio features impact how they are grouped together. The project utilizes PCA in order to reduce the dimensions of the audio feature list before running a K-Means model in order to classify and create custom playlist from the most played tracks. You can access the playlist here:

Twitter is a social media platform that allows users to send and read short messages called "tweets." Tweets are limited to 280 characters or less and can include text, images, and links to other websites. Twitter was founded in 2006 and has since become a popular platform for celebrities, politicians, journalists, and other public figures to share their thoughts and updates with their followers. Twitter has millions of active users and is available in more than 40 languages. With such active user base using the platform to share their opinion, it can be an arduous task to moderate tweets. One such difficulty for twitter is to flag hate tweets. Hate speech is any speech, conduct, writing, or expression that may incite violence or prejudicial action against or by a particular individual or group, or because it disparages or intimidates a particular individual or group. It is generally considered to be harmful and unacceptable, as it can contribute to a culture of intolerance and discrimination. In many countries, hate speech is regulated by law and may be punishable by fines or imprisonment.

However, the definition of hate speech can vary and is often a subject of debate, as it can be difficult to draw a line between protected free speech and harmful expression. Hate speech detection is important for Twitter because it can help the platform to create a safer and more inclusive environment for its users. Hate speech can cause harm to individuals and groups who are targeted, and can also contribute to a culture of intolerance and discrimination. By detecting and removing hate speech, Twitter can help to protect its users from this kind of content and create a more positive and welcoming atmosphere. Additionally, detecting and addressing hate speech can be important for Twitter from a business perspective, as it can help to maintain the platform's reputation and user trust. In recent years, there has been increased pressure on social media platforms to take action against hate speech and other forms of harmful content, and Twitter has implemented a number of policies and tools to help identify and address this type of content on its platform.

This project is like a tip of the tip of the tip of the iceberg (yes, I typed it 3 times.). The project utilises Natural language processing (NLP), which is a field of artificial intelligence that focuses on the interaction between computers and humans through the use of natural language. It involves developing algorithms and models that can understand, interpret, and generate human language. NLP can be used to classify hateful or abusive tweets as part of an effort to combat hate speech on social media platforms. This is still a complex task because labelling tweets as hateful or not is also a big pandora’s box. However, I took a very simple approach which involved cleaning and formatting the raw text data to make it more suitable for analysis. This included tokenization (splitting the text into individual words or symbols), stemming (reducing words to their base form), and removing stop words (common words that do not convey meaning). I classified tweets that either racist or homophobic. I also created a webapp where you can type in a tweet and find out if that tweet is a hate tweet or no. Imagine a much more advanced version of the app can help serve as a guiding light on hate speech.