**Health GSK**

Karenann joined GSK as a Chief Digital & Technology Officer in 2009 with the primary objective of shortening medicine development down to 1 year. This demonstrates how GSK takes its own initiative (Data Analyst Behavioural principle) in using big data and data science/analytics principles to transform its business.

GSK once completed an internal survey on the fragmentation of the data distribution across the businesses, and concluded that a Data Lake is needed to centralise data of different sources to enable greater insight on medicine development. This process highlights GSK has a basic understanding of data life cycle, as well as the importance of data structure and data system design (Data Analyst Knowledge principle). To achieve Data Lake, GSK also enlisted the aid of AI to transform the Data into a ‘clean format’ prior lake ingestion, demonstrating creative thinking whilst ensuring data quality is in its most pristine condition (Data Analyst Behavioural and Knowledge principle).

Perhaps, surprisingly to anyone unfamiliar with the data world, GSK also placed enormous emphasis on data and human interaction. One such example is the mandatory Code of Conduct Training (data-oriented) undertaken by 33,000 employees and the limited access of patience’s’ data by external researchers to ensure data and legal regulations are satisfied (Data Analyst knowledge principle). Besides, GSK also placed customers a great significant as one can choose personal data to be partially or fully available to GSK suggesting the company’s abilities in working with external people in regard to data (Data Analyst Behavioural principle).

**easyJet Data Science opportunity**

After a detailed discussion with a Lead Data Scientist, we identify the below opportunities where one could use Data Science to assist easyJet Operations. First, one could develop a recommendation system to predict absence pattern of a Cabin Crew/Pilot based on their previous work pattern; in turn, easyJet could intervene one’s future work pattern if the system deemed that a particular individual is prone to a high risk of absence. Second, alike to GSK, easyJet has a variety of data formats via different data sources that need to be in production-grade. Therefore, one could use Machine Learning to transform data into production-grade which would enable easyJet to create a more efficient Data product. Third, given a high volume of customers has been requested refunds since the onset of the pandemic, easyJet’s customer service has been overwhelmed by demand. A data solution to such would be building a chatbox based on the integration of Natural Language Processing and Recommendation system.