DARK HORSE CONSULTING GROUP

BRANDON TRAN

Pegasi – Market Forecasting Algorithm

BUSINESS CASE

- Forecasting is an art and an integral tool in the Data Analyst/Scientist toolkit

- As a consultant at DHC, I was pulled into many projects with investors or existing Cell and Gene Therapy (CGT) players who wanted to expand in the industry.

- With Pegasi, DHC aimed to grow away from a simple excel spreadsheet to a simulation algorithm

  - We felt that important market dynamics were not being captured in the excel model

  - We wanted a model informed with real world data

  - We wanted a model that simulated the real process of moving through the clinical trial space

    - Instead of "There were 200 products in Phase 1 this year, so 40% of them (80) move forward to Phase 2"

    - We wanted "Drug product XYZ is entering its 4th year in Phase 1 and sees a XX% success rate compared to product ABC which is entering its 2nd year in Phase 1 which sees a PP% success rate"

WHAT MADE PEGASI DIFFERENT FROM OTHER EXISTING OFFERS

- Real world data
  - Success rates and Duration data gathered from our data intelligence service

- Incorporation of real-world factors not typically modeled
  - Drug products don't move in a straight line down the approval process – Pegasi can model products dropping out of the clinical space

- Individual modeling via MonteCarlo simulation
  - Treating each individual drug as its own entity instead of a bulk/batch group which flows through the pipeline
    - Some products take 2 years while others take 6 years to proceed through a clinical phase

OVERVIEW

- **1) Gather Intelligence**
  - Paid Subscription service

- **2) Clean & Bucket Data**
  - Scrub, bucket, and clean outliers from Intelligence data via Python

- **3) Montecarlo Simulation**
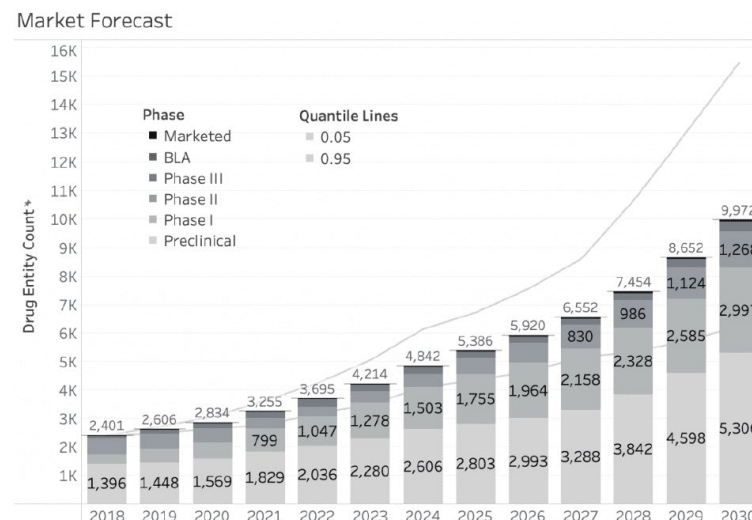 - Run multiple Montecarlo simulations with the cleaned data via Python

- **4) Export Forecast Summary to Excel**
  - Summarize and export results to Excel for Tableau access

- **5) Visualize in Tableau**

Market Forecast

- **6) Analyze/Further Processing**
  - Market Forecast data can be the end-point for investors or companies looking to advance into another field of the CGT industry
  - Market Forecast Data can be used for Facility Build vs Buy models as the demand estimate
    - When combined with COGs, Labor, etc – We were able to provide customers a holistic summary on greenfield vs brownfield decisions.

MODEL VALUE

- When modeled with 2018 clinical trial data
  - Peagsi was run for a 12 year forecast and the estimated clinical trial assets for the year 2020 matched to near perfection with publically available data
    - Phase 1 + Phase 2 entities forecasted = 1001
    - Phase 1 + Phase 2 entities actual count = 1068
    - Phase 3 entities forecasted = 138
    - Phase 3 entities actual count = 152

- This information was used to estimate the current date (2022) market as well as the future growth of the industry in multiple projects with investors and players in the CGT space.