Retention Risk &
Employee Turnover Strategy
Executive Summary: Initial Findings

Prepared for Dr. Donald Wedding, CHRO
Annual HR Summit Day 2
May 12, 2019

People Insights Team
Agnes Cheng
Brent Young
Mary Taylor
Michelle Eure

People Insights

Problem Statement

The Walt Disney Company has undergone significant organizational and leadership changes over the last 5 years which has caused turnover to steadily increase. With the recent 20th Century Fox acquisition and competitive landscape, executives anticipate further disruption leading to the possible loss of high potential and key talent.

Repercussions of high voluntary turnover:

- High replacement costs for the company.
- Projects get delayed in a time where meeting deadlines are crucial.
- Work needs to be redistributed to the remaining employees, which causes worker fatigue, burnout, and employee dissatisfaction.
- New talent has to be recruited, trained and given time to acclimatize themselves to the company.
- Employee satisfaction goes down.
- Loss of institutional knowledge.

Cost Avoidance

Estimated Attrition cost per departing employee: $75,000

Average salary of $50, 000 * 150% cost of turnover

Estimated Cost Avoidance per year with a 10% reduction: $3 million

$75,000 * 400 voluntary exits per year*10% reduction

Note: This is not actual data.

Business Objectives

Objective	Deliverable	Success Criteria
• Identify high, medium, and low flight risk employees.	• Exploratory data analysis of demographics, survey data, and attrition for employees. Potentially exploring cluster analysis. • Build predictive models to predict the probability that someone would leave the company.	• Decrease in voluntary turnover rate and cost avoidance. • Increase in retaining key talent. • Accurate models.
• Determine which factors prompt employees to stay vs. leave.	• Build an explainable predictive model to uncover key insights and themes.	• Changes to the work environment, policies, HR strategy, etc. • Decrease in voluntary turnover rate and costs.

People Insights

		• Increase in employee satisfaction scores.
• Create a Stay Survey and use the model results to help retain high-risk employees.	• Create questions, theoretical experiment process, and A/B testing approach.	• Shift from reactive to proactive mindset amongst leaders and HR. • Increase in retaining key talent and career conversations. • Increase in employee satisfaction scores.
• Create an interactive dashboard and mobile app for Corporate HR and leaders.	• User-friendly dashboard and app that allows users to see low, medium, and high risk by team and individual along with relevant demographics.	• Usability and buy-in from CHRO, leaders, and HR.

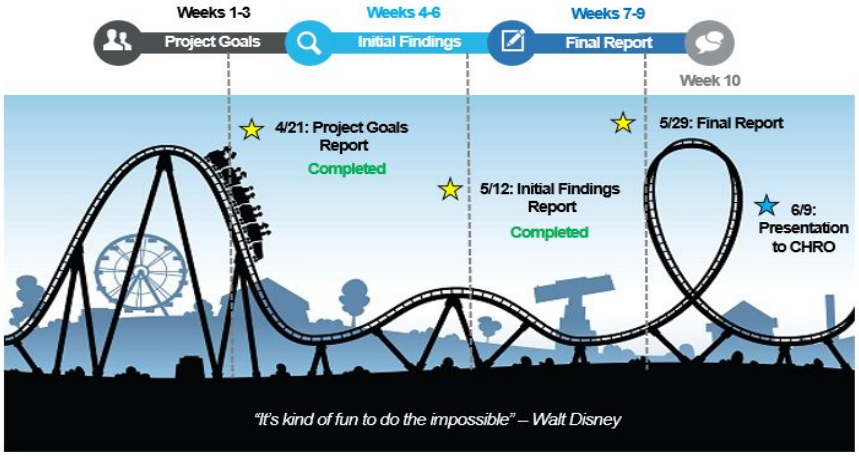
Solution

To help Disney Corporate overcome these challenges, the People Insights' team utilized a data set that contains approximately 1,500 employee records to build several classification models to predict the probability that an employee will leave the company, determine which factors prompt employees to stay vs. leave, and create a Stay Survey so that leaders/HR can use the model results to retain high-risk employees. Overall, this will help inform the development of a Corporate HR strategy to retain top talent, increase employee satisfaction/engagement survey scores, increase productivity, decrease voluntary turnover, and lower hiring/replacement costs, thus saving the company money. An interactive dashboard and mobile interface to share the analysis with Corporate HR and leaders was also created.

Project Status: On Target

Our project has been following an ambitious 10-week project plan and timeline that we've broken up into 3 phases (project goals, initial findings, and final report). The project is progressing on schedule. Our team is productive, meeting at least once or twice a week for 2 -3 hours to consider results, define next steps, and ensure we remain on point. Communication/collaboration has been excellent via WhatsApp, email, and CoLab, the Google code sharing tool. Key deliverables are complete, including

exploratory analysis, feature identification and initial modeling findings. Additional modeling techniques have been identified to refine results and are in progress. The updated timeline is shown below with key milestones identified with stars. We've completed phase 1 and 2 and are now heading into the home stretch.



Key Findings

Observations and Learnings

The best performing model was Logistic Regression. We identified the factors and themes that prompt employees to stay vs. leave the company so that the company can make changes to the work environment, policies, and HR strategy. Here is a list of variables that were included in the model in order of importance.

Observations & Learnings	Why it Makes Intuitive Sense
Employees who have lower total working years are more likely to leave the company.	Employees that have less work experience are usually younger and are eager for advancement and progression.
Employees who logged overtime over the course of 1 year were **7 times** more likely to leave the company compared to those who did not.	Employees may be burned out, overworked, or suffer from fatigue.
Employees who are single are **4 times** more likely to leave the company compared to those who are married (1.6 times).	Single employees are more mobile. They don't have a family to support and/or a wife/husband influencing their decisions to stay or leave a company.

People Insights

Employees who live further away from work are more likely to leave the company.	Cost of gas, inconvenience, and less time with family. Employees may opt for a company that is closer to their home.
Employees that have worked at a number of companies in the past are more likely to leave the company.	One word, “job hoppers”!!!
Employees who are younger are more likely to leave the company.	Employees who are younger are usually more eager for career advancement and progression.
Employees who have low satisfaction survey scores are more likely to leave the company.	Unhappy employees usually leave companies.
Employees who received a low amount of training events in the prior year are more likely to leave the company.	Employees want to be developed and feel appreciated or else employees usually leave.
Employees who have lower daily rates ($) are more likely to leave the company.	Employees usually seek monetary gains and pay raises.
As the tenure of an employee increases, the more likely he/she is to stay with the company.	Higher tenure usually indicates good culture fit, employees that are happy enough to stay, and loyalty to the company.

Data Preparation

- The data was split into a training (70%) and test set (30%).
- The training set will be used to fit the models and the test set will be used to estimate prediction error for model selection.
- K-fold cross-validation, which is a resampling procedure, was also conducted on the training dataset and was used for model selection purposes.
- The best model was then applied to the test set for final evaluation.
- Majority of the variables didn’t include any major outliers.
- Data did not contain any missing values.
- For categorical data, label encoding was applied.

Exploratory Data Analysis

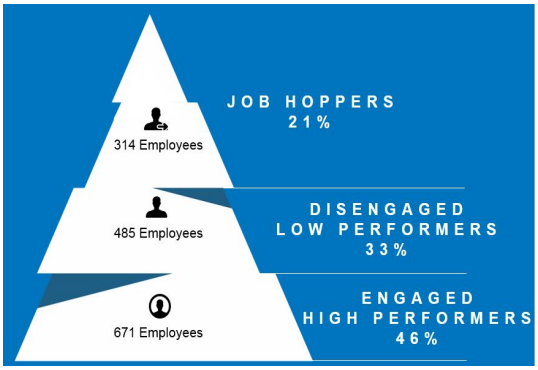
- The data for employees primarily comes from the Research & Development and Sales departments, with HR having the least amount of employees.

People Insights

- Most of the nine job roles are primarily Sales Executives, Research Scientists, and Laboratory Technicians.
- The data set is imbalanced with a low rate of attrition (237 out of 1470 employees left the company) so we are naturally trying to predict “rare events”.
- Majority of the variables were skewed.
- A correlation matrix revealed several variables that were correlated. For instance, Percent Salary Hike vs. Performance Rating.

Feature Selection: The second part of EDA, focused on using bar plots, box plots, and automated variable selection techniques such as univariate feature selection, recursive feature elimination (with logistic regression), tree-based feature selection (feature importance), a decision tree, and lasso embedded method on the training data set. A wide variety of features were identified across each of the options evaluated. *See the observations and learnings section above for the variables that were included.*

Cluster Analysis: The third part of EDA, focused on using cluster analysis (e.g., k-means clustering) to develop clusters using the survey and demographic data so that we can better understand how employees are feeling. This allows HR to develop targeted strategies and policies to increase employee engagement and satisfaction, which can ultimately improve employee retention. Three segments were determined. *See initial findings report for in-depth profiles on each segment. Note: R was used for this analysis.*

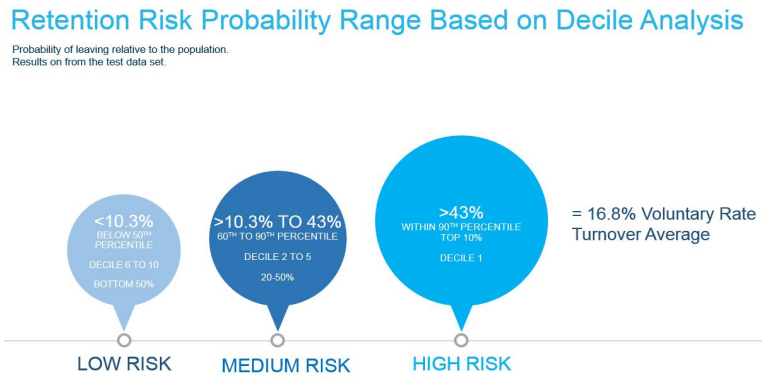


Modeling

Given that this a classification problem, a mix of modeling methods were used; from explainable modeling techniques such as logistic regression and decision trees to more complex modeling techniques such as linear discriminant analysis, XGBoost, SVM, random forest, K-Nearest Neighbors and Neural Network. All modeling was completed in Python and the code was shared using CoLab. The models were assessed using statistical evaluation criteria such as AUC, ROC Curves, Logarithmic loss (Logloss), mean k-folds cross-validation accuracy scores, and Decile Analysis. The Logistic Regression model yielded the best performance metrics which was great news considering it is the most interpretable, simple, and easy to explain. As a result, we decided to move forward with the Logistic Regression model.

People Insights

To determine probability cutoffs for high, medium, and low-risk employees, the People Insights team decided to use the decile analysis on the test data set for the logistic regression model. Here is a visual summary of the retention risk probability range based on the decile analysis:



Way Ahead

Our primary goals in the next 3 weeks are to finalize the modeling approach, harden the feature selection and create the dashboard/mobile app that can be shared with Corporate HR and business leaders. We also plan on continuing to improve our models using resampling techniques such as SMOTE, experiment with feature engineering, flush out the report, conduct additional analyses (time permitting), and create our final presentation. Recommendations to our CHRO will also be made on how the business can implement the model using a Stay Survey approach. Additionally, programs, policy changes, and initiatives based on our model insights will also be recommended.

As a result, we are “on track” to completing our initial goals that were laid out at the beginning of this project. However, if something changes in the next 3 weeks, we will let the CHRO know in our status reports. It’s been an honor to work with the Disney team and contribute to the Disney of the future. We look forward to sharing our final results in the upcoming weeks.