Next Best Action Lead Reason Analysis - An Attribution Approach

Presented by Charlie & Francis
Owned by the Business Analytics Team at Fidelity Canada

RoadMap

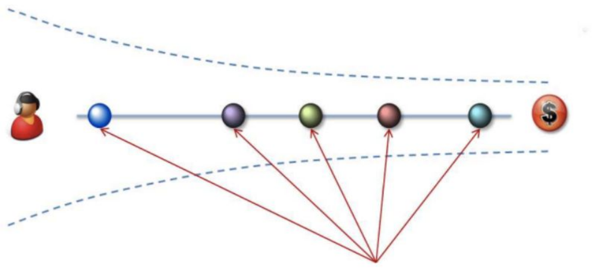
- Introduction
  - Project Overview
  - Background Information
  - Model Explanation & Selection
  - Preliminary Model Results (Francis)
- Data Preprocessing
  - Exploratory Data Analysis
  - Decision-Making about Journey Aggregation
- Results & Discussions
  - Final Thoughts & Summary

Introduction

This section will introduce the **project background**, define **technical jargons** and explain different **attribution models** used.

Project Overview

Aim & Scope: How different closed lead reasons impact customers' purchase behavior - measured by an actual big gross purchase ticket conversion (Value > $10K).



Assign credit % to each lead reason

Metrics: Each closed lead reason will have a corresponding contribution percentage towards if there is any large gross purchase.

Success Criteria: One/Several successful models generating NBA lead reason attributions.

Centre for Management of Technology & Entrepreneurship

Jargons Explanation

NBA: Next Best Action for the sales team to take on.

Closed Leads/Lead Reasons: Leading Reasons for why a customer contacting Fidelity Canada and is followed up by a sales representative.

Conversion: Customer’s Single Gross Purchasing Value > $10K.

Customer’s Purchasing Path: A path showing each customer’s purchasing behaviour over the study period.

Conversion Path: A subpath which finally lead to a conversion.

Null Path: A subpath which doesn’t lead to a conversion.

Total Conversion Values per Path: Total conversion value summed up per purchasing subpath.

Jargons Explanation

Attribution Models: A set of rules to determine how credits for *conversions* is assigned to each closed lead reasons path.

Gross Purchase Label: Binary Classifier (1/0) indicating if there is a big ticket conversion (> $10k).

Heuristic Attribution Models: Static methods which computes each channels (leads) attribution towards a final gross purchase using a deterministic approach.

Data-Driven Attribution Models: Using Data Science/Machine Learning methods to calculate each channels (leads) performance towards GP.

NBA Lead Reasons

- **Service Standard / BDM Service Standard:** Fidelity's Service Standard.
- **Redemption Risk:** Risks from mutual funds being early redeemed.
- **Trending Low in Seasons GS:** Low Trend in Seasonal Gross Sales.
- **Seasons Prospect:** Seasonal GS Forecast.
- **Fund Reco:** Fund Recommendations.
- **Web Leads:** Advisor Website fund Viewing Activities.
- **Webcast Reco:** Webcast Recommendations.
- **Webcast Follow Up:** Follow Up through webcast.
- **Short-Term +ve GS Momentum:** Average of Recent Purchase History Going Up
- **Short-Term -ve GS Momentum:** Average of Recent Purchase History Going Down
- **High Fund/PM Concentration:** High Concentration
- **Recommendation List:** Other Reasons

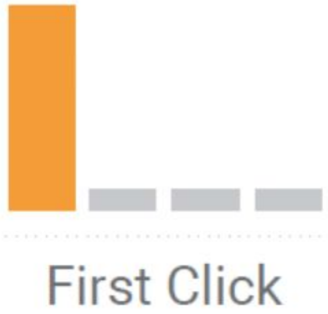
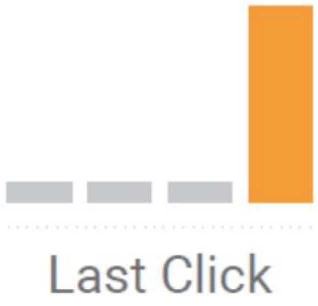
Background Information

Study Subject: Closed NBA Leads since *2021-01-01* - the sales team has ***accepted and made*** the contact, following NBA lead outputs.

Study Scope: 10 Months (YTD).

Models Selection: Heuristic Models & Statistical/ML Models.

Model Explanation - Single Touch Attribution Models

First-Touch Attribution Model	**Last-Touch** Attribution Model
Assumptions: First lead reason in user's path gets ***all*** contribution values (100% credit) to GP.	**Assumptions: Last** lead reason in user's path gets ***all*** contribution values (100% credit) to GP.
	

Model Explanation - Multi-Touch Attribution Models

Linear Attribution Model

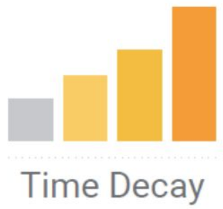
Assumptions: **All** lead reasons in the user’s path ***equally share*** the contribution value towards GP.

Position-Based Attribution Model

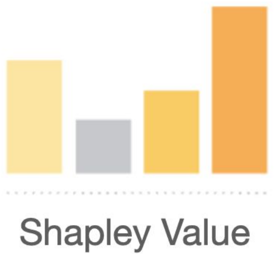
Assumptions: **First and Last** lead reason in the user’s path gets ***40%*** contribution values each to GP; the **remaining 20%** being ***equally divided*** into the remaining lead reasons.

Time-Decay Attribution Model

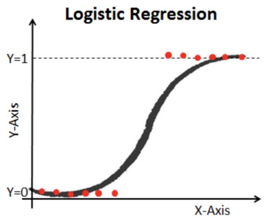
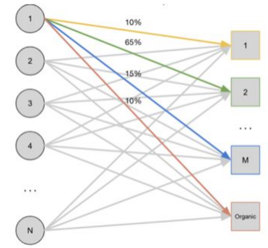
Assumptions: the **later lead reasons** on the user’s path get ***more*** contribution value towards GP. Last lead reason gets the ***most*** credits.



Model Explanation - Data-Driven Attribution Models

The **Shapley Value** Attribution Model	The **Markov Chain** Attribution Model
<u>**Assumptions:**</u> Each lead reason work **together** in order to drive conversions - Contribution made by each lead is calculated using the Shapley Value (A Game Theory Approach).	<u>**Assumptions:**</u> Using **Transition Matrix** to determine the *possibilities* of each IP's path from the starting lead to the final lead.
	

Model Explanation - Data-Driven Attribution Models

The **Bagged Logistic Regression** Attribution Model	The **Probabilistic** Attribution Model
<u>Assumptions:</u> A **stable and predictive** logistic regression approach to calculate each leads contribution towards conversion. Randomly sampling IPs from the raw dataset for 2000 times, bagging results together to update the logistic regression coefficients.	**<u>Assumptions:</u>** Using **Second-Order Probability Estimation** (Probability Theory) to generate each lead's probability towards a conversion. Assuming each reason is working independently.
	

Probabilistic Model for Channel Attribution

CMTE UNIVERSITY OF TORONTO

Centre for Management of Technology & Entrepreneurship

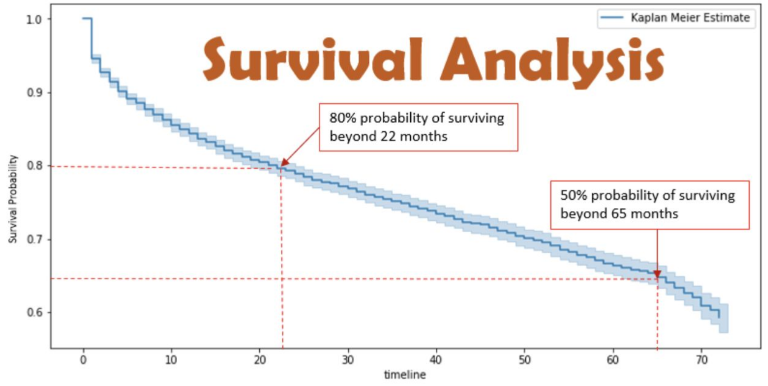
Model Explanation - Data-Driven Attribution Models

The **Additive Hazard** Attribution Model

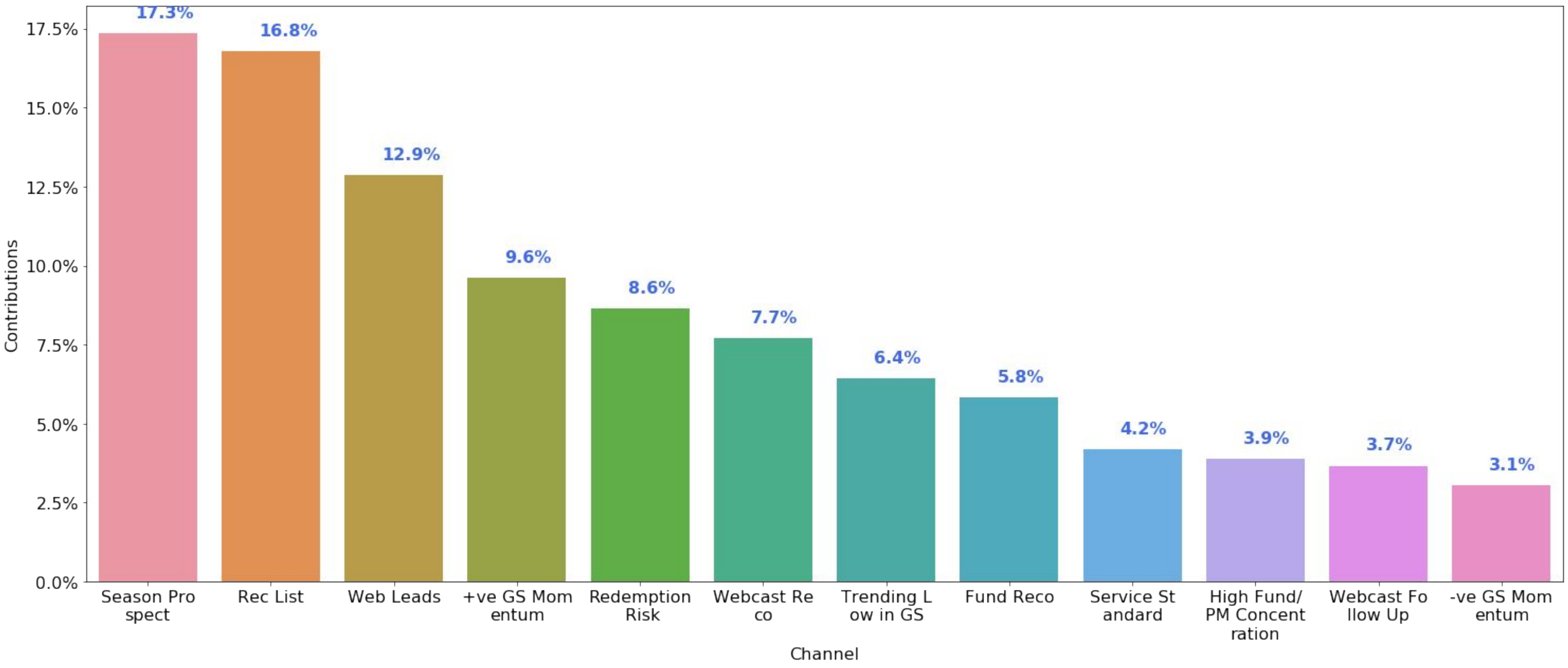
<u>**Assumptions:**</u> Using **Survival Theory** analyzing both the *variations of the time-decay speed* and the *impact strength* of each channel.

Additive Hazard Model: Uses the concept of survival theory to calculate the **probability** of a conversion for each chanel (lead reasons) and use it as the contribution of each channel.

Advantage: Able to *Predict* future conversions and *removes the presentation bias.*



Normalized Shapley Value (90 days Journey) Results



Data Preprocessing

This section I will discuss about how I formed the customer's **purchasing path**, **exploratory data analysis** and **Data Cleaning**

Data Description

User Path: A user's purchasing path - could lead to a ***conversion or null***.

Sample Path:

Seasons Prospect -> Fund Reco -> Webcast Reco -> Short-Term Positive GS Momentum -> Redemption Risk -> *Conversion*

OR

Seasons Prospect -> Fund Reco -> Webcast Reco -> Short-Term Positive GS Momentum -> Redemption Risk -> *Null*

Data Cleaning

Path Aggregation: Aggregate each IP's purchasing path by considering all Lead reasons available.

Raw Dataset:

ip_id	created_da	Service Sta	Redempti	Trending L	RRSP/SPRI	Fund Reco	Web Lead	Webcast F	Webcast F	Short Terr	Short Terr	High Fund	Rec List	GP
108327	3/27/2021	0	0	0	0	0	0	0	0	0	1	0	0	0
108388	5/14/2021	1	0	0	0	0	0	0	0	0	0	0	0	0
108471	6/3/2021	0	0	0	1	1	1	0	0	0	0	0	0	0
108471	9/10/2021	0	0	0	1	1	0	0	0	0	0	0	0	0
108471	9/25/2021	0	0	0	1	1	0	0	0	0	0	0	0	0
108480	3/27/2021	1	0	0	1	1	0	0	0	0	0	0	0	0
108480	6/4/2021	0	1	0	1	1	0	1	0	0	1	0	0	0
108480	7/15/2021	1	1	1	0	1	0	1	0	0	0	0	0	0
108480	8/13/2021	0	1	0	1	1	0	1	0	1	0	0	0	0
108480	8/31/2021	0	1	0	1	1	0	0	0	0	0	0	0	1

Sample User Purchasing Path

Example: ID - 108480

Potential Path:

- Service Standard -> NULL
- Service Standard -> Redemption Risk -> NULL
- Service Standard -> Redemption Risk -> Seasons Prospect -> NULL
- Service Standard -> Redemption Risk -> Seasons Prospect -> Webcast Reco -> NULL
- Service Standard -> Redemption Risk -> Seasons Prospect -> Webcast Reco -> Fund Reco -> Conversion

Sample Purchasing Journey - With GP

User ID: 108480 = 1125 Possible Permutations!

3 x 5 x 5 x 5 x 3

3

- Service Standard
- Seasons Prospect
- Fund Reco

5

- Redemption Risk
- Seasons Prospect
- Fund Reco
- Webcast Reco
- Short Term -ve GS Momentum

5

- Service Standard
- Redemption Risk
- Trending Low in RRSP GS
- Fund Reco
- Webcast Reco

5

- Redemption Risk
- Seasons Prospect
- Fund Reco
- Webcast Reco
- Short Term +ve GS Momentum

3

- Redemption Risk
- Seasons Prospect
- Fund Reco

Sample Purchasing Journey - W/O GP

User ID: 108480 = 182250 Possible Permutations!

3 x 3 x 5 x 3 x 5 x 3 x ... x 2

3
- Service Standard
- Seasons Prospect
- Fund Reco

3
- Null
- Conversion
- Continue

5
- Redemption Risk
- Seasons Prospect
- Fund Reco
- Webcast Reco
- Short Term -ve GS Momentum

3
- Null
- Conversion
- Continue

5
- Service Standard
- Redemption Risk
- Trending Low in RRSP GS
- Fund Reco
- Webcast Reco

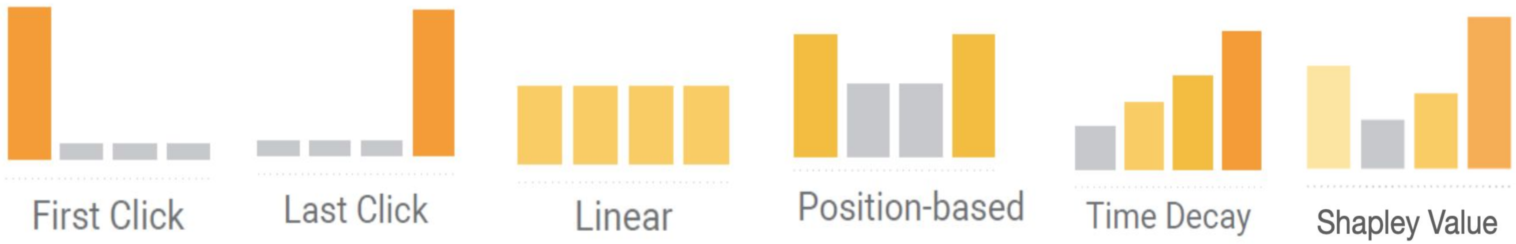
3
- Null
- Conversion
- Continue

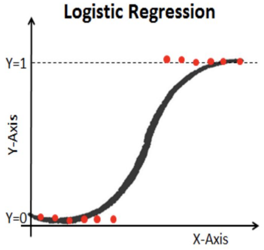
2
- Null
- Conversion

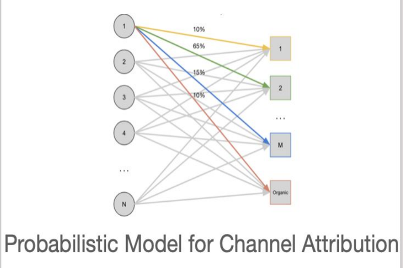
Model Explanation on a Single Journey

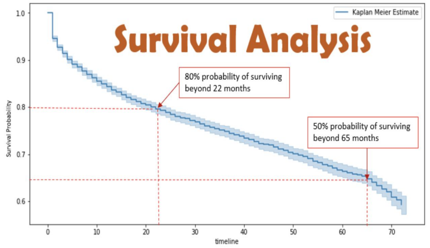
User ID: 108480

- **Sample Journey:** Service Standard -> Redemption Risk -> Seasons Prospect -> Webcast Reco -> Fund Reco -> GP









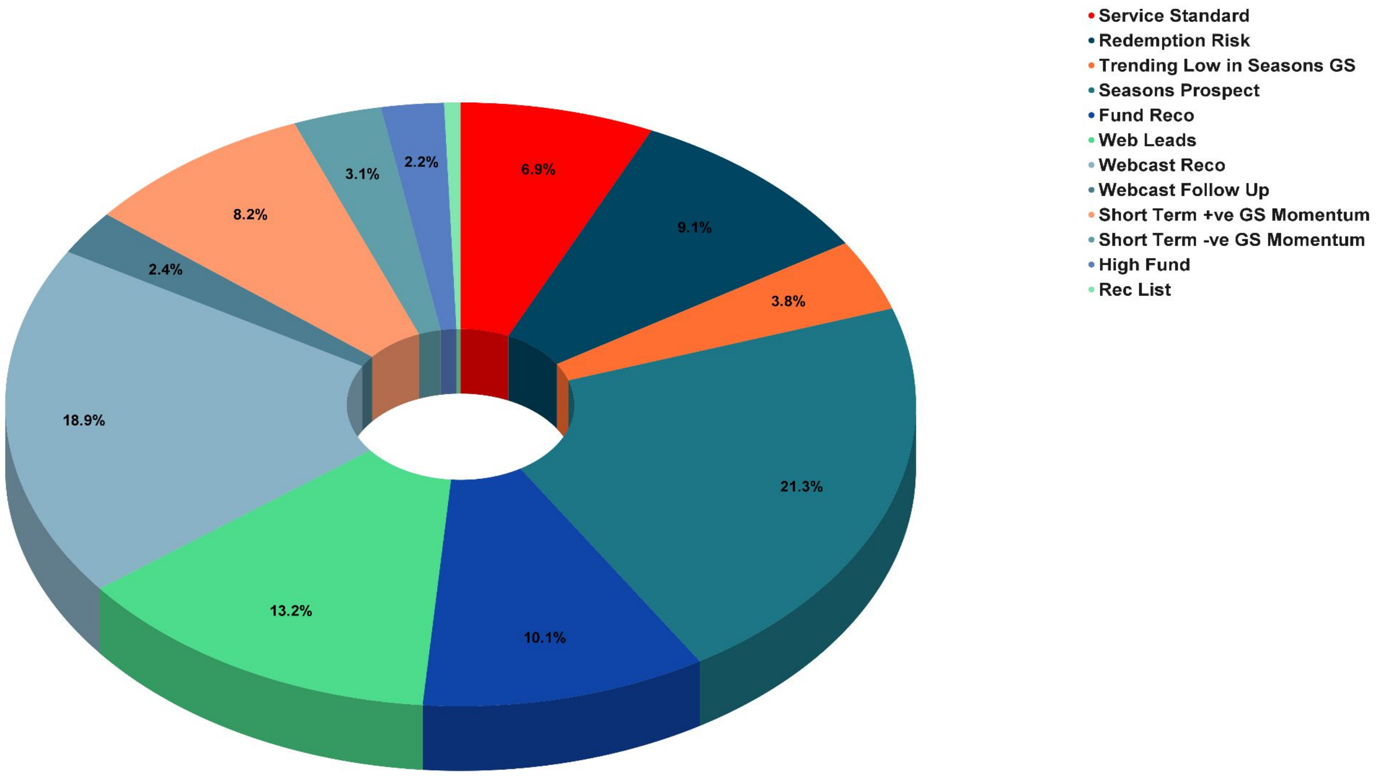
Exploratory Data Analysis

Top 10 Paths Leading to a Conversion:

- Seasons Prospect → GP
- Fund Reco → GP
- Seasons Prospect → Seasons Prospect → GP
- Fund Reco → Fund Reco → GP
- Webcast Follow Up → GP
- Seasons Prospect → Fund Reco → GP
- Fund Reco → Seasons Prospect → GP
- Web Leads → Web Leads → GP
- Webcast Reco → GP
- Webcast Follow Up → Webcast Follow Up → GP

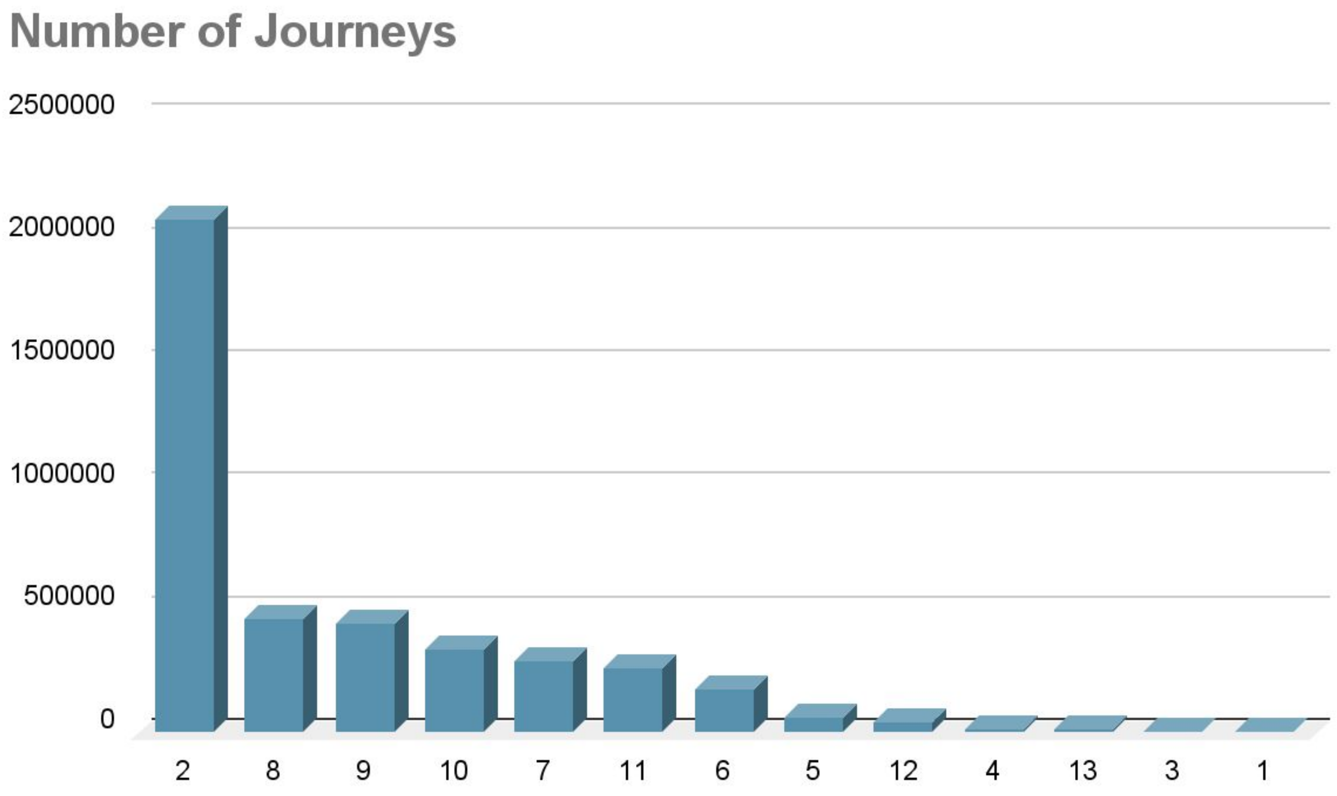
Exploratory Data Analysis

Percentages of All Reasons Leading to a Conversion:



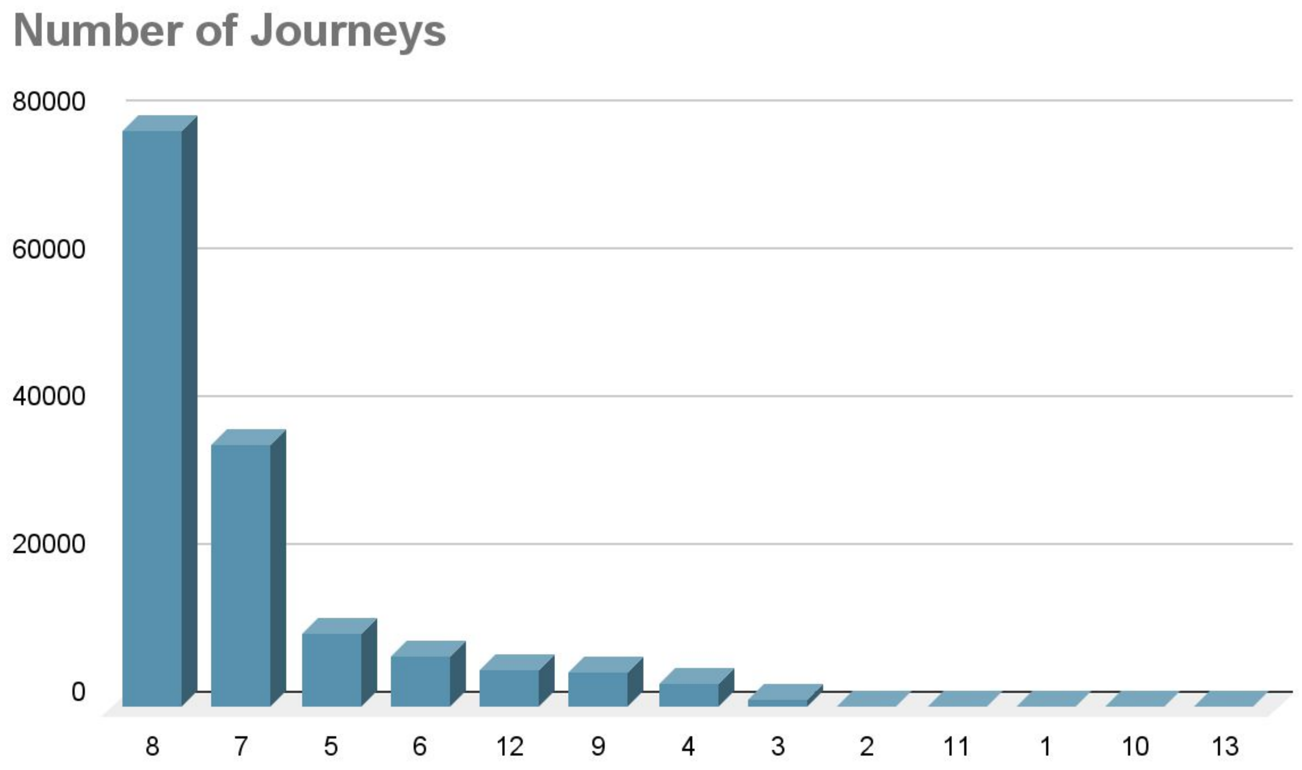
Exploratory Data Analysis

Journey Length Frequencies



Exploratory Data Analysis

Journey (with GP) Length Frequencies



Results and Summary

In this section, I will present the result of each of the 10 models and a discussion on model selection

Top 3 Lead Reasons (Before Normalization)

Model Name	First Lead Reason	Second Lead Reason	Third Lead Reason
First-Touch	Web Leads	Webcast Follow Up	Seasons Prospect
Last-Touch	Seasons Prospect	Fund Reco	Web Leads
Linear	Seasons Prospect	Fund Reco	Webcast Follow Up
U-Shaped (40% First & Last, 20% all others)	Seasons Prospect	Fund Reco	Webcast Follow Up
Time Decay	Seasons Prospect	Fund Reco	Webcast Follow Up
Shapley Value	Seasons Prospect	Fund Reco	Webcast Follow Up
Markov Chain	Seasons Prospect	Webcast Follow Up	Fund Reco
Logistic Regression	Trending Low in Seasonal GS	High Fund/PM Concentration	Service Standard
Probabilistic	Webcast Reco	Seasons Prospect	Fund Reco
Additive Hazard (Survival Theory)	Fund Reco	Webcast Reco	Seasons Prospect

Top 3 Lead Reasons (After Normalization)

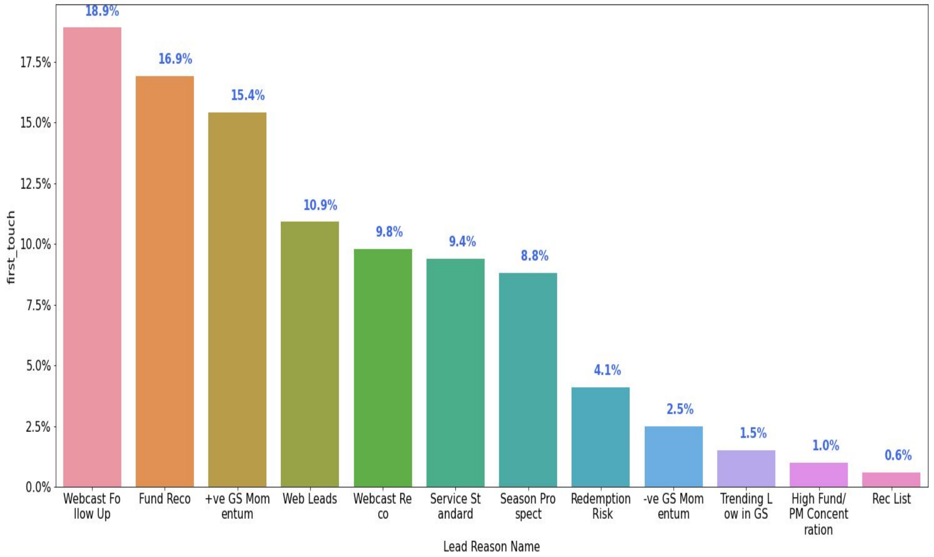
Model Name	First Lead Reason	Second Lead Reason	Third Lead Reason
First-Touch	Webcast Follow Up	Fund Reco	+ve GS Momentum
Last-Touch	Short Term +ve GS Momentum	Seasons Prospect	Webcast Reco
Linear	Seasons Prospect	+ve GS Momentum	Web Leads
U-Shaped (40% First & Last, 20% all others)	Seasons Prospect	+ve GS Momentum	Web Leads
Time Decay	Seasons Prospect	+ve GS Momentum	Webcast Reco
Shapley Value	Short Term +ve GS Momentum	Rec List	Webcast Reco
Markov Chain	Short Term -ve GS Momentum	Seasons Prospect	Webcast Reco
Logistic Regression	Redemption Risk	-ve GS Momentum	Trending Low in GS
Probabilistic	Short Term +ve GS Momentum	Rec List	Seasons Prospect
Additive Hazard (Survival Theory)	Short Term +ve GS Momentum	Seasons Prospect	Webcast Reco

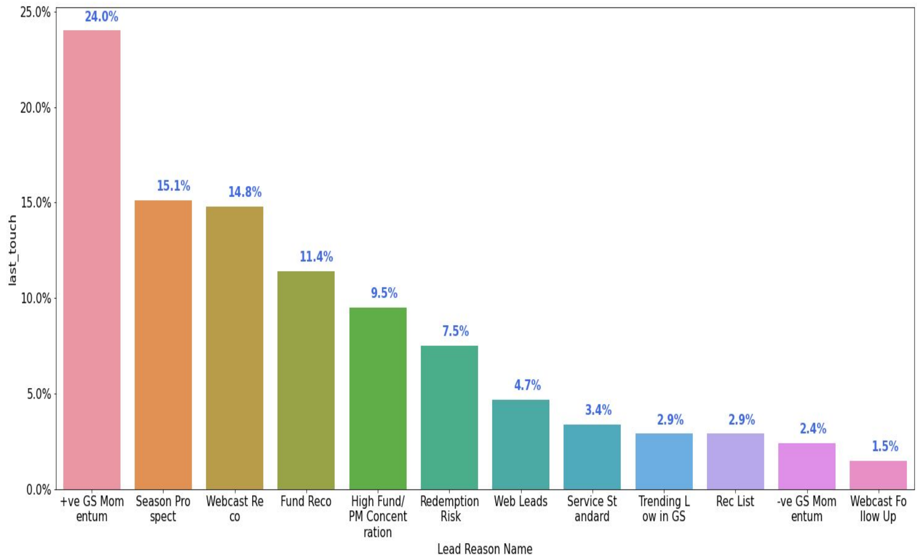
- Normalized Results using the Frequencies of each Lead Reasons

Heuristic Models Results & Discussion

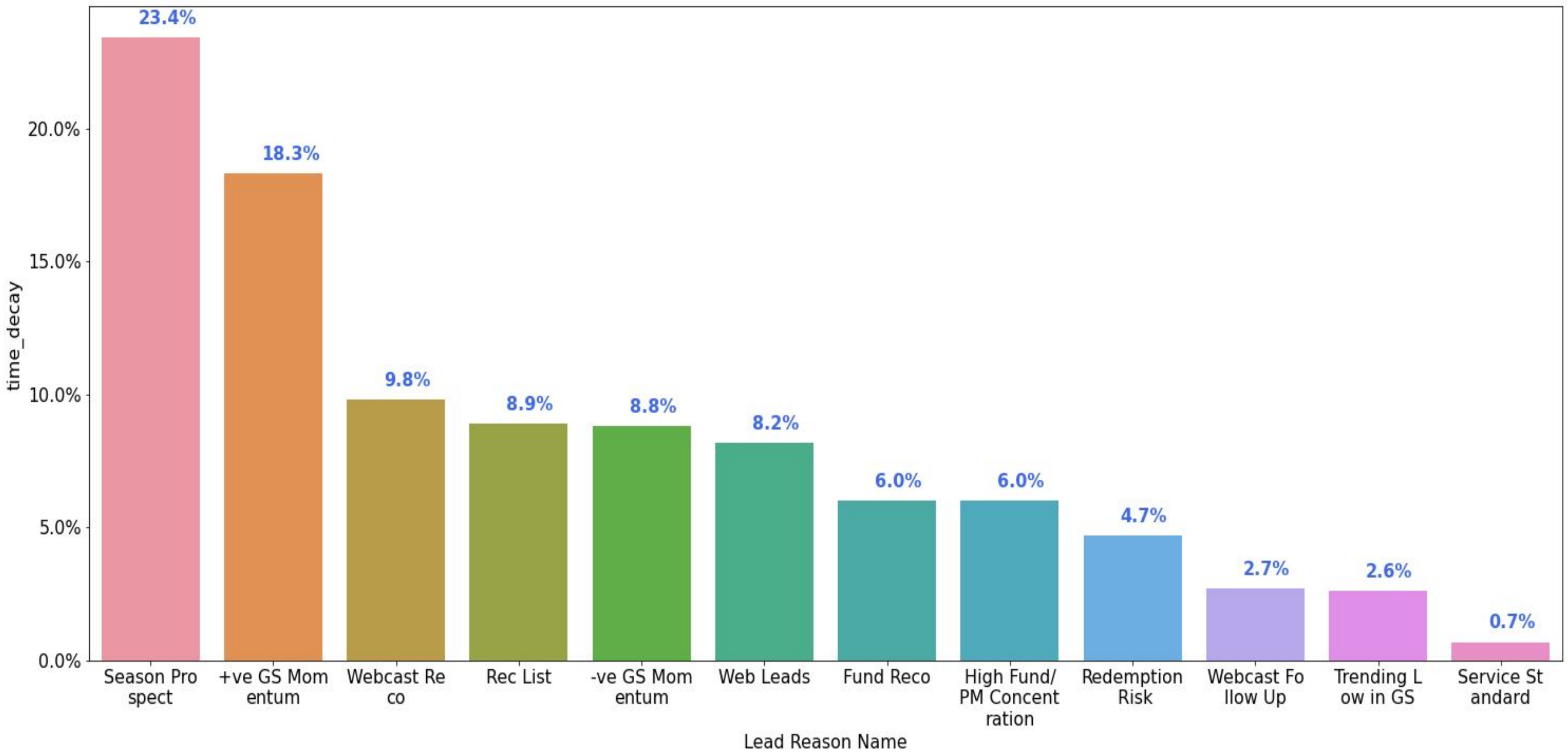
Single Touch Attribution Model Results

First-Touch Attribution Model	**Last-Touch** Attribution Model
<u>**Assumptions:**</u> **First** lead reason in user's path gets ***all*** contribution values (100% credit) to GP.	<u>**Assumptions:**</u> **Last** lead reason in user's path gets ***all*** contribution values (100% credit) to GP.





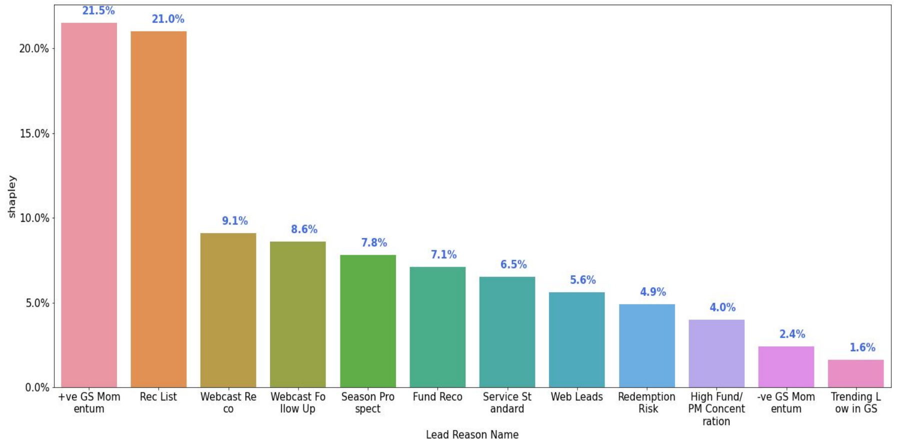
Multi - Touch Attribution Model Results

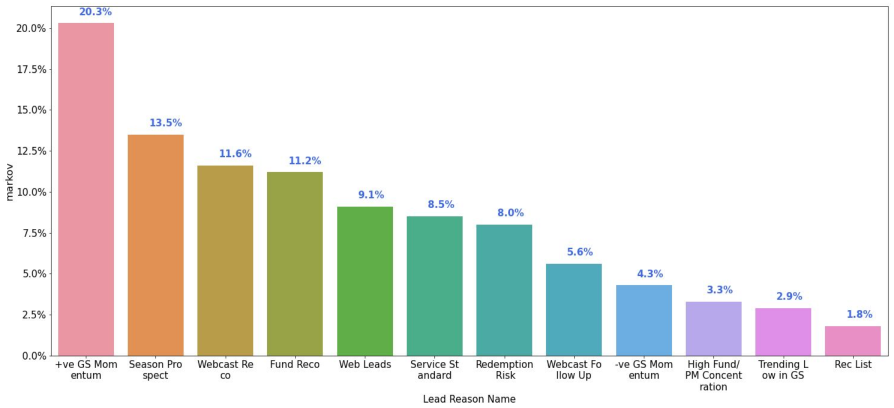


Data-Driven Statistical/ML Model Results

Model Explanation - Data-Driven Attribution Models

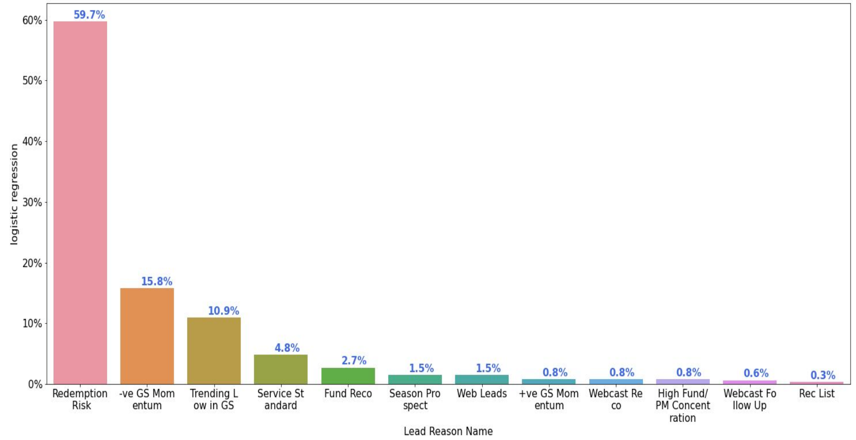
The **Shapley Value** Attribution Model	The **Markov Chain** Attribution Model
<u>**Assumptions:**</u> Each lead reason work **together** in order to drive conversions - Contribution made by each lead is calculated using the Shapley Value (A Game Theory Approach).	<u>**Assumptions:**</u> Using **Transition Matrix** to determine the *possibilities* of each IP's path from the starting lead to the final lead.

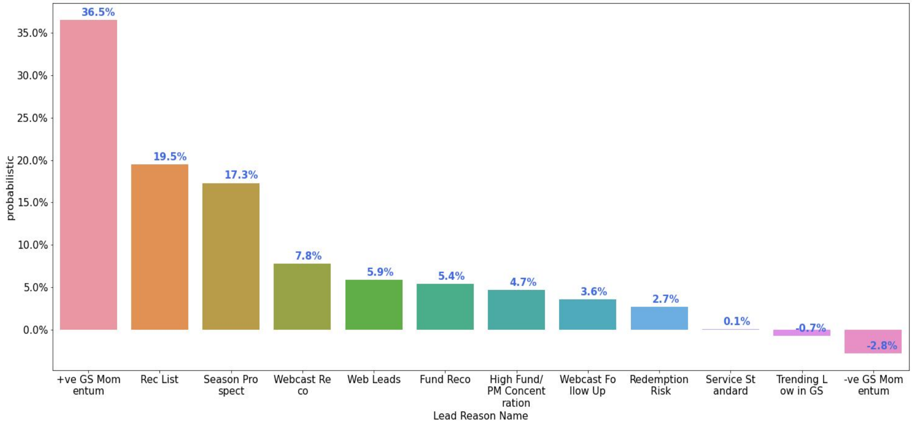




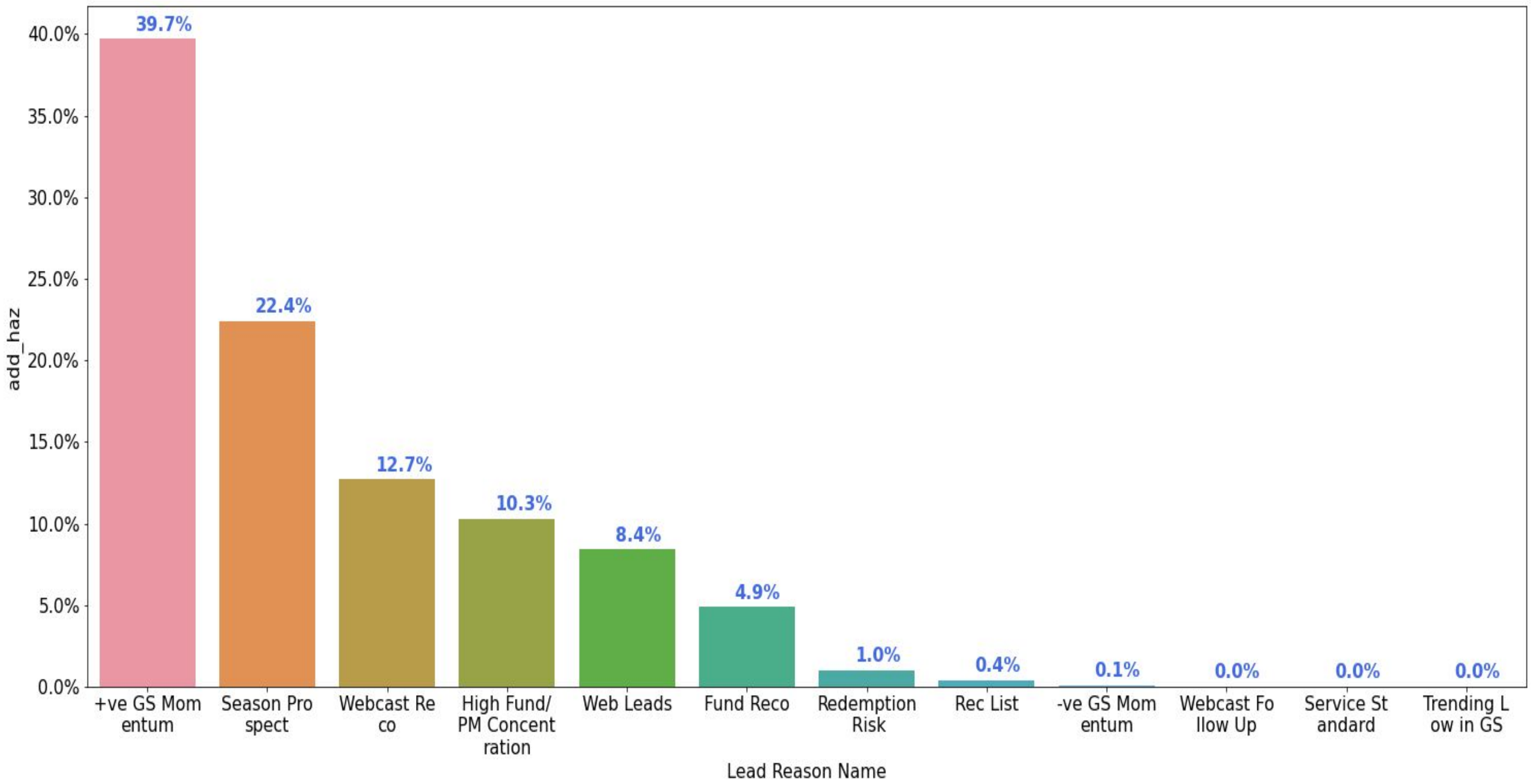
Model Explanation - Data-Driven Attribution Models

The **Bagged Logistic Regression** Attribution Model	The **Probabilistic** Attribution Model
Assumptions: A **stable and predictive** logistic regression approach to calculate each leads contribution towards conversion.	**Assumptions:** Using **Second-Order Probability Estimation** (Probability Theory) to generate each lead's contribution towards conversion.





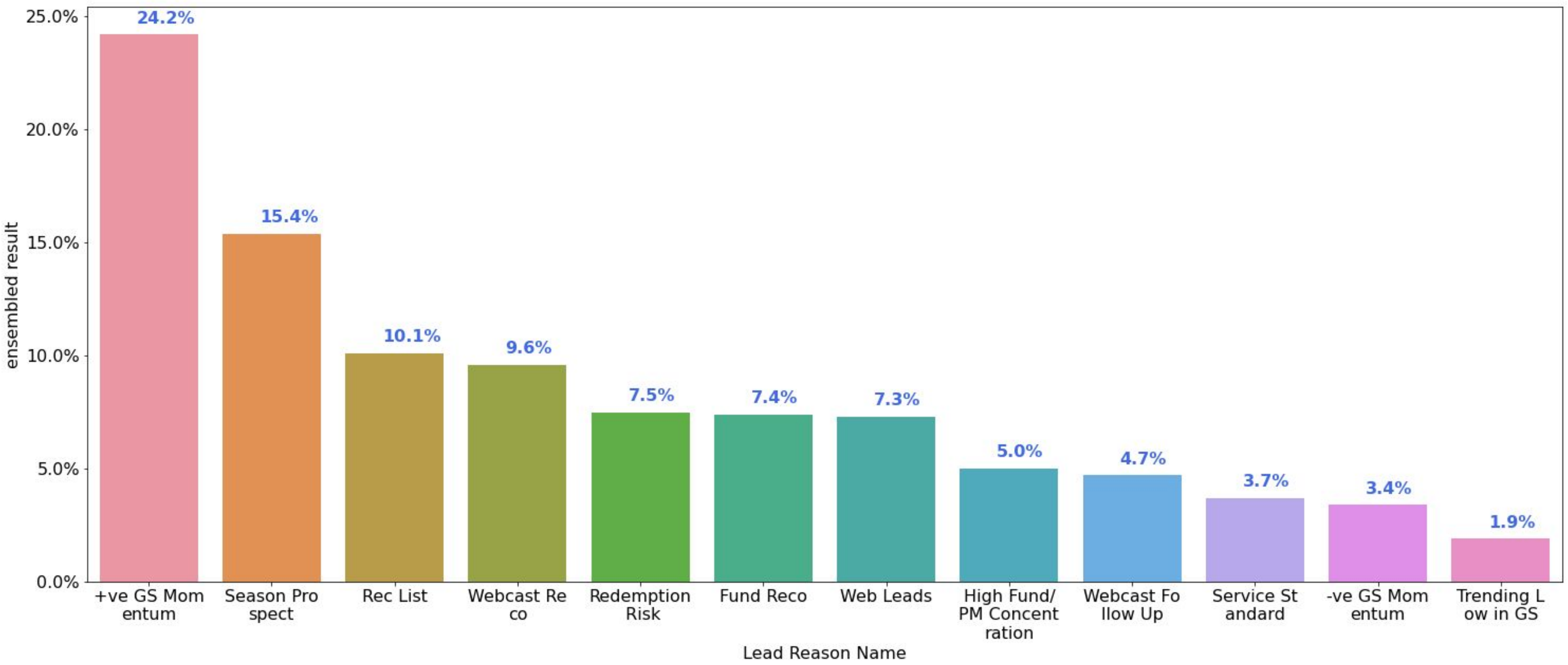
Model Explanation - Data-Driven Attribution Models



Final Thoughts, Discussions & Results

- **Time-Decay Model** is the best among Heuristic Models
- Shapley Value model fails to consider sequence of data
- Logistic Regression model requires independence among the reasons, which lead to weird results
- Probabilistic Model does not consider how each reason works together to get the conversions and often have negative results.
- Additive hazard model takes too long to train and often get skewed results
- **Markov Chain** model is the best among data-driven models
- **Last-Touch, Linear, Position-Based, Time-Decay, Shapley Value, Markov-Chain & Probabilistic models** are recommended.
- Results should consider the assembly of the above-recommended models

Ensembled Results



Final Thoughts, Discussions & Results

- Top 3 Reasons to Proceed for the marketing team: **Short Term +ve GS Momentum, Seasons Prospect & Webcast Follow Up**
- Marketing team should focus on the following order for the leads: **Short Term +ve GS Momentum > Seasons Prospect > Rec List > Webcast Reco > Redemption Risk > Fund Reco > Web Leads > High Fund > Webcast Follow Up > Service Standard > -ve GS Momentum> Trending Low in Seasonal GS**
- Potential Next Steps for the Attribution Models:
  - Research some other data-driven attribution models (LSTM-RNN-Attention Model)
  - Use neural networks
  - Compared the results with models selected in this project
- Our models can be generalized to any other attribution problems

Thank You!

Appendix

This section includes all model assumptions and their corresponding pros & cons

Model Explanation - Single Touch Attribution Models

First-Touch Attribution Model	
<u>**Assumptions:**</u> **First** lead reason in user's path gets ***all*** contribution values (100% credit) to GP.	
Pros ● Simple to set up. ● Easy to implement . ● Great to use if no journey available in the user's purchasing path.	**Cons** ● Doesn't consider each IP's purchasing journey. ● Overemphasize on top-of-funnel activities (leads). ● Doesn't consider how each lead reasons work together to contribute to a final gross purchase.

Model Explanation - Single Touch Attribution Models

Last-Touch Attribution Model	
<u>**Assumptions:**</u> **Last** lead reason in user's path gets ***all*** contribution values (100% credit) to GP.	
Pros	**Cons**
● Simple to set up. ● Easy to implement. ● Great at increasing conversion rates. ● Great to use if no journeys in the user's path.	● Doesn't consider each IP's purchasing journey. ● Overemphasize on the last lead. ● Doesn't consider how each lead reasons work together to contribute to a final gross purchase.

Model Explanation - Multi-Touch Attribution Models

Linear Attribution Model	
Assumptions: **All** lead reasons in the user's path ***equally share*** the contribution value towards GP.	
Pros	**Cons**
● Takes a Multi-Touch Approach ● Easy to implement & get the results ● Great to use if there's a journey in the user's path.	● Model is too simple. ● Treats every lead reasons the same way. ● Ignores the difference between each lead reasons and doesn't acknowledge how they cooperate together to lead to a gross purchase.

Model Explanation - Multi-Touch Attribution Models

<table>
<tr><th colspan="2">Position-Based Attribution Model</th></tr>
<tr><td colspan="2"><u>Assumptions:</u> First and Last lead reason in the user’s path gets 40% contribution values each to GP; the remaining 20% being equally divided into the remaining lead reasons.</td></tr>
<tr><td>Pros
● Takes a Multi-Touch Approach
● Give more credits to more important lead reasons (i.e first and last)
● Improved upon the linear model by weighting the contributions of each lead reason differently.</td><td>Cons
● Model is still simple and restrictive.
● Middle lead reasons might be underestimated.
● Doesn’t acknowledge how they cooperate together to lead to a gross purchase.</td></tr>
</table>

Model Explanation - Multi-Touch Attribution Models

Time-Decay Attribution Model

<u>Assumptions:</u> the **later lead reasons** on the user's path get ***more*** contribution value towards GP. Last lead reason gets the ***most*** credits.

Pros	Cons
● Takes a Multi-Touch Approach ● Improved upon the position-based model by re-weighting the contributions of each lead reasons differently based on their position w.r.t the conversion.	● Assign credits using a linear function to each lead reason. (e.g $y = 2*x/7$) ● Leads might not work in that way in reality. ● Still doesn't acknowledge how they cooperate together to lead to a gross purchase.

Model Explanation - Data-Driven Attribution Models

Shapley Value Attribution Model	
<u>Assumptions:</u> Each lead reasons ***works together*** as a game to the final gross purchase. The leads with the ***highest shapley value*** gets the most credits.	
Pros ● Considers all reasons and how they cooperate together to lead to a gross purchase ● Improved upon the single-touched and multi-touch attribution models by providing a way to fairly dividing the payoff between the players.	**Cons** ● Does not consider the sequence of purchasing ● Does not consider the occurrence counts of each lead reason

Model Explanation - Data-Driven Attribution Models

Markov Chain Attribution Model

Assumptions: Each lead reason is a ***Markov State*** with the probabilities (learnt using our dataset) to go to the ***next state***.

Pros	**Cons & Improvements**
• Improved upon the shapley value approach by considering the sequence of purchasing • The transition probabilities of each state is only dependent on the previous state.	• Transition matrix could not be calculated if there’s no journeys in the path. • Could be better if we use bayesian methods in estimating the transition matrix.

Model Explanation - Data-Driven Attribution Models

Bagged Logistic Regression Attribution Model	
Assumptions: Randomly sampling a batch (**25%** of the total size) from the original dataset and run logistic regression on it. Repeat the process for **2000 times**.	
Pros ● Use a regression approach to classify conversions ● Accurately updates the coefficient by repeating the process for a large amount of time. ● Simple and easy to implement.	**Cons & Improvements** ● Unable to capture non-linearities ● Overfitting ● Requires each independent variable have no correlations with each other. ● Does not consider how the reasons work together to achieve the final conversion.

Model Explanation - Data-Driven Attribution Models

<table>
<tr><th colspan="2">Probabilistic Attribution Model</th></tr>
<tr><td colspan="2"><u>Assumptions:</u> Each lead reason as a channel could have a probability lead to a conversion. By calculating the second order probabilities of the conversion, one could get the total credit of each lead reason.</td></tr>
<tr><td>Pros

● Improved upon the markov chain approach by considering the probabilities of each channel leading to a conversion.</td><td>Cons & Improvements

● Does not consider how the reasons work together to get the final conversion.
● Does not consider the sequence of the input data.
● Often get negative results</td></tr>
</table>

Model Explanation - Data-Driven Attribution Models

Additive Hazard Attribution Model	
Assumptions: Treat each channel's conversion path as a survival rate. As time goes by, the survival rate will decrease and thus has less impact on final conversion. It uses additive hazard rate to reflect the influence of relative channels on the conversion.	
Pros ● Improved upon the Probabilistic Model by also considering the time-decay model into channel attributions using survival theory.	**Cons & Improvements** ● Training takes too long ● Does not consider contextual information and intrinsic user conversion rate into account ● Skewed Results