SQL Design - Carvana Recommendation Engine

Background

The mission statement of Carvana, an online car purchasing service, "is to change the way people buy cars by providing our customers a car-buying solution that is fun, fast, fair, and powered by technology." To help Carvana achieve this goal, we decided to create a SQL powered recommendation system to assist consumers with the car-buying process.

Currently, customers are unable to make informed decisions regarding which type of car to purchase (can only filter on color, mileage, brand, etc.). Our database plan solves this issue by allowing customers to input information regarding their personalities and spending habits to recommend vehicles matching their preferences. This system will increase Carvana's revenue by targeting customers through recommendations (much like Amazon does) while educating our customers on the types of cars that will best fit their needs. The business requirements involve combining customer information - obtained by users signing up for an account through Facebook or the Carvana sign-in portal and submitting a survey containing personality and budgeting information - and browsing data (e.g., cookies) along with inventory levels to create recommendations for customers.

Process Changes

As outlined in our vision above, the proposed changes documented in our TO-BE flow will allow for the creation of a robust recommendation system. This system will help to improve our lead conversion rate by increasing the clarity and ease of the car buying process which will help to increase inventory turnover, market share, and cash flow.

Specific process changes to the user experience include: the recommendation prompt site users receive when first visiting carvana.com, the creation/inclusion of a finance/personality survey during user account creation, the use of cookies to help feed the recommendation system, and the recommendation engine itself. Each of the changes will be implemented primarily to feed data into the recommendation engine to increase the accuracy of the system.

Gap Analysis

The gaps in the AS-IS process are addressed by the colored segments in the TO-BE flow. The most notable gap in the existing (AS-IS) process is the lack of a recommendation engine. In the TO-BE diagram this was addressed by adding an additional swimlane. Our trigger event of our process is that the customer chose the option on the webpage that requires personalized

recommendations, and our inputs include the basic information, personal aesthetics, personality traits, social habits, credit score, etc. Our output will be a proper recommendation to the customer that suggests suitable cars to them. Accordingly, in our TO-BE swimlane, additional changes were made to accommodate the various prompts/information required to allow the recommendation to function (e.g. "Need a recommendation," "Have an account," account type, personality/finance survey, cookies, etc.).

Queries and Reports

Through our ten sample tables, we generated several queries to reveal patterns in the data.

1. As a car recommendation system, we want to see customers that have an annual income of > $100,000, and we may recommend high-end cars.

```sql
select a.FirstName || ' ' || a.LastName as "Full Name",
a.CarvanaID, s.AnnualIncome
from personalityorfinancesurvey s join CarvanaAccount a
on a.CarvanaID = s.CarvanaID
where s.AnnualIncome>100000;
```

The results that the following six customers have an annual income higher than $100,000. These customers may be able to afford higher-end cars.

	Full Name	CARVANAID	ANNUALINCOME
1	Jared Bergantino	1120345	100500
2	Wang Fang	1947553	123055
3	Li Guo	3302506	3000000
4	Phillis Wheatley	3925967	7004304
5	Bob Dylan	7557293	4950302
6	Melinda Gates	7664821	390000

2. We generate our recommendations based on multiple indexes including the survey. We want to see what colors of car customers with different PersonalityTraits would like.

```sql
select AestheticPreferences as "Preferred Color",
listagg (PersonalityTraits,',') within group (order by PersonalityTraits) as "Traits",
listagg (CarvanaID,',') within group (order by CarvanaID) AS "CarvanaID"
from personalityorfinancesurvey
group by AestheticPreferences;
```

The results show us that people with different personal traits may have different preferences for the color of cars. For example, outgoing persons may prefer red cars, and people pursuing honesty and purity prefer white more.

	Preferred Color	Traits	CarvanaID
1	Black	Sophisticated,mystery	1120345,3925967
2	Blue	optimistic	7664821
3	Brown	thrifty	7557293
4	Gold	humourous	6491739
5	Gray	Cautious	3302506
6	Red	Outgoing	1947553
7	Silver	Practical	8604934
8	White	honesty,purity	418561,8606313

3. When we have recommendations, we need to make sure we have such cars in stock. Choose the cars in stock to match the personal traits of customers based on the results from the survey.

```
select s.PersonalityTraits as "Traits", i.StockNumber, i.CarModel, i.bodytype
from personalityorfinancesurvey s join recommendation r
on s.SurveyID = r.SurveyID
join Recommendation_Inventory ri on r.RecommendationID = ri.RecommendationID
join Inventory i on i.StockNumber = ri.StockNumber;
```

The result shows us the traits of customers and the corresponding car (including the stock number, the car model, and the body type) that will be recommended to that customer.

	Traits	STOCKNUMBER	CARMODEL	BODYTYPE
1	Sophisticated	14234	323	Hatchback
2	Cautious	95002	S Class	Sedan
3	purity	59403	Panamera	Sport
4	Outgoing	50402	Integra	Hatchback
5	Practical	30494	S Series	Sedan
6	humourous	10043	Jetta	Sedan
7	honesty	19403	Jetta	Hatchback
8	optimistic	14235	CR-V	Crossover
9	mystery	12634	5 Series	Sedan
10	thrifty	12434	Highlander	SUV

4. Knowing that some customers sign in with Facebook accounts, we want to know the relationship between the number of likes on Facebook and traits and Aesthetic Preferences of customers.

```
select f.likes, a.FirstName, s.AestheticPreferences, s.PersonalityTraits
from facebooklogin f join CarvanaAccount a
on f.FCarvanaID = a.CarvanaID
join personalityorfinancesurvey s
on a.CarvanaID = s.CarvanaID
order by f.likes;
```

The results can show the relationship between the total number of likes with the personal traits.

	LIKES	FIRSTNAME	AESTHETICPREFERENCES	PERSONALITYTRAITS
1	123	Stephen	Silver	Practical
2	442	Bob	Brown	thrifty
3	2232	Phillis	Black	mystery
4	3232	Ai	Gold	humourous
5	33232	Melinda	Blue	optimistic
6	43434	Fan	White	honesty

5. To allow a customer to buy a car, he or she has to have high enough credit scores, we want to see whether customers with higher income tend to have higher credit scores.

```
select c.CreditScore, a.CarvanaID, S.AnnualIncome
FROM Credit c join  CarvanaAccount a
on c.CreditID = a.CreditID
join personalityorfinancesurvey s
on a.CarvanaID = s.CarvanaID
order by 1,3;
```

The results give the relationship between the annual income and the credit score.

	CREDITSCORE	CARVANAID	ANNUALINCOME
1	245	8604934	80400
2	433	8606313	49583
3	556	6491739	65000
4	556	3925967	7004304
5	601	7557293	4950302
6	632	7664821	390000
7	682	1947553	123055
8	720	1120345	100500
9	773	418561	39503
10	840	3302506	3000000

6. Based on our current system, we want to see which cars are recommended to which customer.

```
SELECT a.CarvanaID "Carvana ID", a.LastName "Last Name", a.FirstName "First Name",
EXTRACT(YEAR FROM i.caryear)||' '||i.Make||' '||i.CarModel "Recommendation"
FROM CarvanaAccount a
JOIN PersonalityOrFinanceSurvey s ON a.CarvanaID = s.CarvanaID
JOIN Recommendation r ON s.SurveyID = r.SurveyID
JOIN Recommendation_Inventory ri ON r.RecommendationID = ri.RecommendationID
JOIN Inventory i ON ri.StockNumber = i.StockNumber
ORDER BY 2;
```

The results show the match of the customers and recommendations in our current system.

	Carvana ID	Last Name	First Name	Recommendation
1	1120345	Bergantino	Jared	1986 Mazda 323
2	418561	Bingbing	Fan	2001 Volkswagen Jetta
3	7557293	Dylan	Bob	2005 Toyota Highlander
4	1947553	Fang	Wang	1990 Acura Integra
5	7664821	Gates	Melinda	2012 Honda CR-V
6	3302506	Guo	Li	1996 Mercedes Benz S Class
7	8604934	King	Stephen	2001 Saturn S Series
8	6491739	Weiwei	Ai	2007 Volkswagen Jetta
9	3925967	Wheatley	Phillis	2006 BMW 5 Series
10	8606313	Zhao	Ziyi	2011 Porsche Panamera

7. We set the 670 as the threshold of the credit score to buy a car. We want to see which customers are approved to purchase a vehicle based on credit.

```
SELECT a.carvanaid, a.lastname, a.firstname,
    CASE WHEN (c.creditscore > 670) THEN 'Approved'
    ELSE 'Not Approved' END "CREDIT_CHECK"
FROM Credit c
JOIN CarvanaAccount a ON c.CreditID = a.CreditID
ORDER BY 4,2;
```

The results show whether the customers are allowed to buy a car based on the credit check.

	CARVANAID	LASTNAME	FIRSTNAME	CREDIT_CHECK
1	1120345	Bergantino	Jared	Approved
2	418561	Bingbing	Fan	Approved
3	1947553	Fang	Wang	Approved
4	3302506	Guo	Li	Approved
5	7557293	Dylan	Bob	Not Approved
6	7664821	Gates	Melinda	Not Approved
7	8604934	King	Stephen	Not Approved
8	6491739	Weiwei	Ai	Not Approved
9	3925967	Wheatley	Phillis	Not Approved
10	8606313	Zhao	Ziyi	Not Approved

8. We want to see if there is a relationship between the address of a customer and drivetrain (awd, fwd, rwd)

```
SELECT a.carvanaid,a.address, i.drivertrain
from carvanaaccount a
join PersonalityOrFinanceSurvey s on a.CarvanaID = s.CarvanaID
JOIN Recommendation r ON s.SurveyID = r.SurveyID
JOIN Recommendation_Inventory ri ON r.RecommendationID = ri.RecommendationID
JOIN Inventory i ON ri.StockNumber = i.StockNumber
;
```

The results show the address of customers and their corresponding drivetrain.

	CARVANAID	ADDRESS	DRIVERTRAIN
1	1120345	36 Park St Hartford CT	FWD
2	3302506	4 Main St Avon CT	RWD
3	8606313	63 Mulberry Ave New York NY	AWD
4	1947553	947 Fifth St Denver CO	FWD
5	8604934	3 South St Des Moines IA	RWD
6	6491739	573 Wild Blvd Helena MT	FWD
7	418561	45 Bridge Rd Tallahassee FL	FWD
8	7664821	3 Q Ave Atlanta GA	AWD
9	3925967	1039 Red Rd Bismark ND	FWD
10	7557293	643 Holly Dr Las Vegas NV	4WD

9. We also want to see if the personality traits of customers will affect their choice toward the make of car.

```sql
Select s.PersonalityTraits, i.Make
from personalityorfinancesurvey s
JOIN Recommendation r ON s.SurveyID = r.SurveyID
JOIN Recommendation_Inventory ri ON r.RecommendationID = ri.RecommendationID
JOIN Inventory i ON ri.StockNumber = i.StockNumber
;
```

The results show the relationship between personal traits and the make of the car.

	PERSONALITYTRAITS	MAKE
1	Sophisticated	Mazda
2	Cautious	Mercedes Benz
3	purity	Porsche
4	Outgoing	Acura
5	Practical	Saturn
6	humourous	Volkswagen
7	honesty	Volkswagen
8	optimistic	Honda
9	mystery	BMW
10	thrifty	Toyota

Business value

The most important idea of our recommendation system is to utilize the information in the personality or finance survey to generate personal car recommendations. We can research the annual income, Personality Traits, and Aesthetic Preferences of customers to determine cars recommended in terms of brands, colors, models, body types, etc. Customers can be more engaged in the personalized selection process, and the company can benefit from the following perspectives.

I Customer

From the customer's point of view, customers can buy a car that suits them more easily, instead of having to search for specific types of cars through their own research. Especially for customers who have little knowledge about vehicles, it is hard for them to know the make, function etc. of a specific vehicle, while it is easy for them to judge if this car is suitable for themselves when you present a vehicle right in front of them. The closer our recommendations meet their expectation, the more likely they will be to buy a car on Carvana. Therefore, this is a more efficient process, and customer satisfaction will increase as a result of the better shopping experience.

II Profitability and development

The expected conversion rate will increase from 2.0% to 5.0%, and the expected inventory turnover will increase from 1.4% to 2.0%. From the perspective of profitability and company development, by using the recommendation system, customers have more opportunities to choose cars that are more in line with their personal preferences and aesthetics, in other words, the recommendation system is "smarter" than the past system. Therefore, the conversion rate and the inventory turnover will increase, so the transaction volume and revenue of our website will increase. Accordingly, the company's profits will improve, which will support the company's long-term development and growth.

III Market share

The expected market share will increase from 0.4% to 1.0%. From the perspective of market share and industry competition, if our competitors still use traditional methods, but we apply the car recommendation systems through innovation, in addition to improving profitability and user satisfaction, we can further increase market competitiveness and occupy more market share. From a long-term perspective, our innovative recommendation system can help the company continue to expand and eventually become a market giant.

Conclusion

The rise of Carvana provides customers with a new way to purchase cars. Our goal is to build a "smart" recommendation system to generate recommendations for customers to assist them. We focus on the personality, financial situation, and aesthetic preferences to generate a recommendation system that helps customers find cars so we can increase our profit as well as occupy more market share and gain advantages in competition with other companies.