Report for VDP Pilot Performance Analysis Project

Date of Report

24 October 2023

Author

Zeng Fanyi

zfygni@hotmail.com

Data Driven | Business Insights | Master of NUS

1. Introduction

1.1 Project Introduction

Our group project was centred around a digital transformation consultancy engagement with an automotive company, which is detailed below.

Considering data privacy and permissions issues, Huayang Aotong, the dealer, lacks access to customer profile data from Audi China and its joint ventures, FAW Audi. Audi China aims to address this by establishing a data platform that integrates customer profiles from all three parties and provides them to the dealers. This integration will empower dealers like Huayang Aotong with more opportunities to connect with potential customers, ultimately driving profit growth.

Our current focus contains the construction and pilot implementation of the Vehicle Data Platform (VDP) for after-sales maintenance in vehicle after-sales services. We are also actively fine-tuning the VDP to augment its functionalities, enhancing its ability to support maintenance services effectively. The primary objective of this project is to unveil the untapped potential of Huayang Aotong in the vehicle maintenance business.

This project involves key stakeholders, including Audi China, the joint ventures, and Huayang Aotong, collectively referred to as the "Three Parties," who are data providers for our VDP project.

Table 1-1 Introduction to Main Subjects

Subject Names	***Abbreviation***	***Remark***
Vehicle Data Platform	VDP	Products and deliverable of the main project
Audi China	Audi China	Our clients and project sponsors
Joint Ventures	JV	FAW Audi, vehicle manufacturer
Huayang Aotong	Hua Ao	Vehicle dealers, the main target of the VDP pilot
Three Parties	Three Parties	Audi China, JV, and Hua Ao
Business Lead	Lead	Refers to vehicle and customer data provided by VDP
Vehicle Aftermarket	Vehicle Aftermarket	After-sales service for vehicles
First Maintenance	First Maintenance	First maintenance after vehicle purchase
Regular Maintenance	Regular Maintenance	Regular maintenance of vehicles

1.2 Industry Overview

China's automotive landscape has witnessed a pronounced slowdown after enjoying a period of robust growth. The industry, which once rocketed from a mere 2.31 million units in 2003 to a staggering 21.15 million units in 2018, has hit some turbulence. Post-2018, the market's momentum took a noticeable downturn, hinting at an industry-wide saturation. By 2023, projections indicate a subtle drop to 20.31 million units, emphasizing the market's cooling pace.

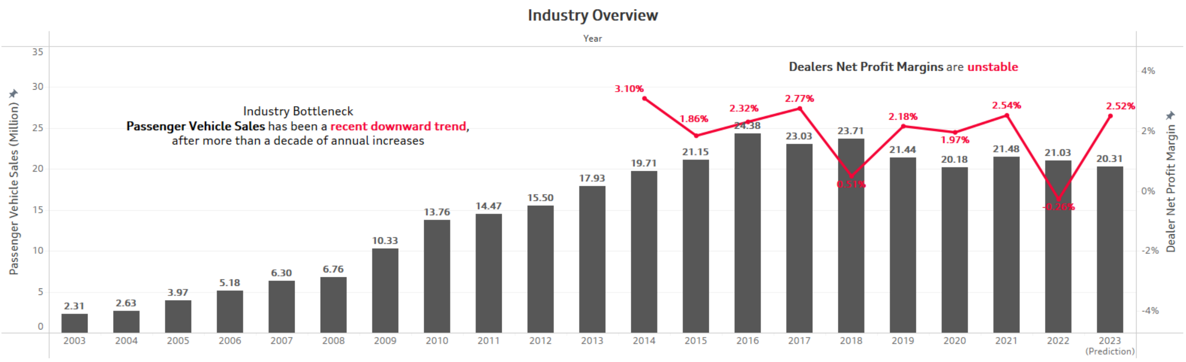

Figure 1-1 Industry Overview

Coupled with this subdued sales growth is the volatile nature of dealers' net profit margins. The metrics paint a compelling story of the industry's challenges: a high of 3.10% in 2013 gave way to a worrisome 0.26% in 2020. These dramatic swings in profitability shed light on the vulnerabilities of the dealership ecosystem.

These shifting dynamics and financial unpredictability underscore the critical need for industry players to diversify their revenue streams. Amid such ebbs and flows, it becomes imperative for automotive dealers to recalibrate their focus towards after-sales services, hunting for fresh avenues of growth. In particular, the post-sale car maintenance segment emerges as a promising frontier. Drawing from these market observations, we spotlight our project's central subject, "Hua Ao." As a downstream affiliate of Audi, Hua Ao not only retails Audi automobiles but also specializes in post-purchase maintenance. Yet, challenges loom large, as they grapple with hurdles in amplifying their maintenance service offerings. Our expertise aims to navigate them past these impediments.

1.3 Business Problem

Currently, there are two primary challenges facing Hua Ao in their maintenance business:

1) Insufficient Lead Volume:

One of the pressing challenges facing Hua Ao dealership revolves around the limited availability of leads at their disposal. In this context, a 'lead' pertains to detailed information about vehicle owners who may soon require car maintenance services. Currently, Hua Ao's database primarily consists of individuals who have purchased vehicles directly from their establishment. To establish a connection and promote their maintenance services, the dealership's primary method of communication is through telephone interviews. Unfortunately, a significant portion of this potential customer information is stored with the joint venture (FAW Audi) and Audi China, placing most of these leads beyond Hua Ao's reach. This limitation has a ripple effect since the dedicated team at Hua Ao, responsible for telephone-based outreach, finds themselves handicapped, unable to reach out to this segment of potential customers. The result is a missed opportunity to engage potential clients, ultimately impeding the dealership's ability to optimize its revenue streams.

2) Inefficiency in Tele-Interviews:

Hua Ao's tele-interview methodology involved unsystematically selecting vehicle owners from their client database. The lack of a systematic approach has impeded their ability to connect with potential patrons effectively. This random selection diminishes the potential impact and efficiency of their outreach efforts, leading to missed opportunities for revenue. It's crucial for Hua Ao to re-evaluate and adjust their strategy to ensure each call is strategically placed, maximizing the chance of a favourable outcome.

1.4 Project Objective

1) Business Objective:

① Obtain more leads from other departments:

The objective is to broaden the customer base for maintenance services, including not only Hua Ao's direct buyers but also potential clientele associated with the joint venture and Audi China, with a primary focus on acquiring additional leads. This expansion aims to create new revenue streams.

② Replace unsystematic selection with rule-based selection:

To better target potential vehicle owners for maintenance services at the dealership, we are moving away from a scattergun approach. Instead, we will adopt a systematic method, utilizing specific rules to determine who to contact. By making more targeted calls, we aim to improve the effectiveness of our outreach and enhance our success rate.

2) Technical Objective:

① Construct VDP (Vehicle data platform):

Develop an integrated database called vehicle data platform (VDP) that merges customer leads from various sources, including Hua Ao, Audi China, and the joint venture (FAW Audi).

② Build selection model to prioritize potential clients:

We utilize algorithms, business insight, and advanced data analysis methods to identify the key determinants influencing customers' maintenance plans. This approach allows us to build models that pinpoint customers more likely to avail maintenance services. By discerning patterns in customer behaviour, we can create targeted outreach strategies, prioritizing contact with potential clientele.

1.5 Project Design

Summarizing the above project objectives and solutions, the Project Design process is obtained as follows.

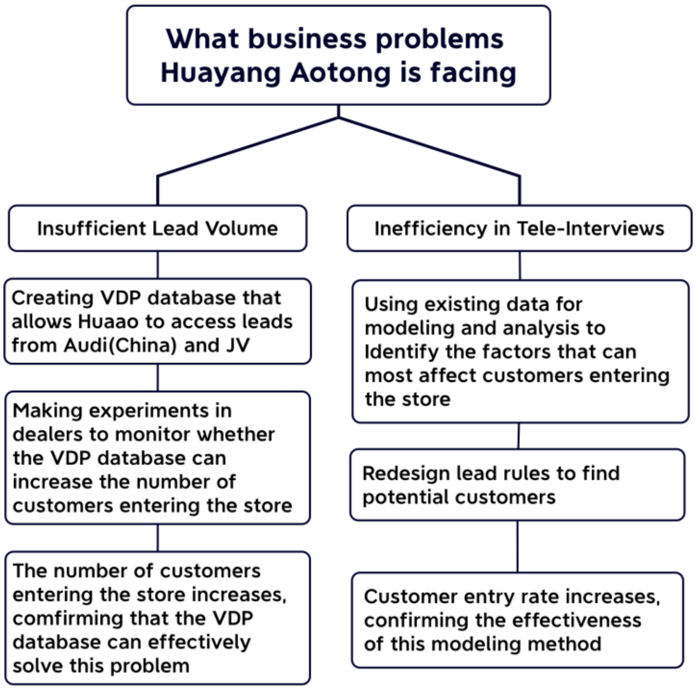

Figure 1-2 Project Design

2. VDP Data Overview

2.1 VDP Data Sources

This section describes the data sources for the VDP system.

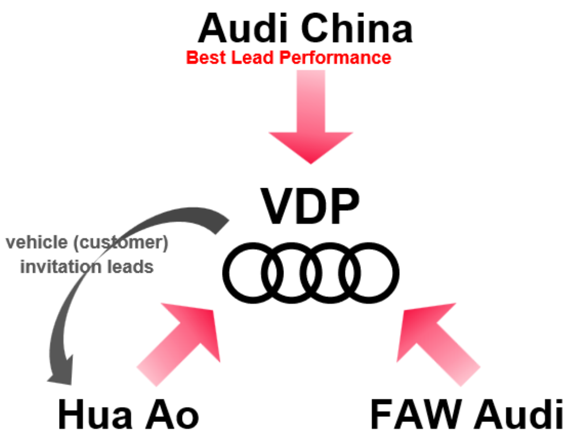

Figure 2-1 Data Sources-Three Parties

Our Vehicle Data Platform has three data sources in total, which are Audi China, FAW Audi, and Hua Ao, the so-called “three parties” (As shown in Figure 2-1). These three parties provide data such as vehicle information, customer information, and in-store maintenance information.

After collecting vehicle related data from the three parties, the VDP system generates a list of vehicles (customer) invitation leads and assigns this leads list to Audi dealers like Hua Ao for them to contact customers for their vehicle maintenance.

Through early-stage lead quality tracking in project, we have found that Audi China has a best leads performance in vehicle invitation when it comes to in-store maintenance. Therefore, in the VDP pilot phase, we integrated data from Audi China to generate the list of vehicle invitation leads and analysed customer in-store performance.

2.2 Data Processing

Before conducting specific analysis, we need to process the data within the VDP system. This section will describe the data preprocessing methods used in the project.

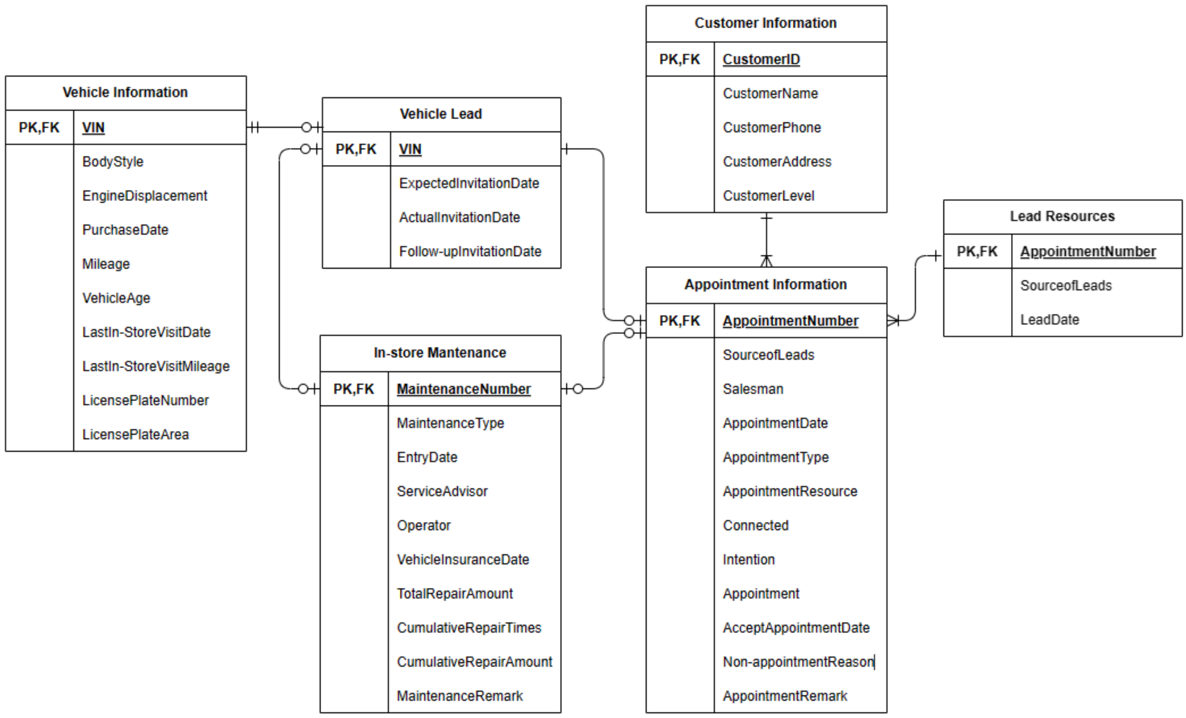

Figure 2-2 VDP Data ERD

Figure 2-2 illustrates the entity relationship diagram for VDP data. It should be noted that in the pilot project, the in-store data includes not only customers who made appointments from the vehicle lead list but also those who entered the store without an appointment. Therefore, there is a 0/1 to 0/1 relationship in the ERD.

1) Data Merging

The data merging process in the Vehicle Data Platform involves using the Vehicle Unique Identification Number ('VIN' code) as the primary key to consolidate several tables. Additionally, other key fields such as 'Maintenance Number', 'Customer ID', and 'Appointment Number' are used for this process.

In this project, we use the VLOOKUP function of Excel to merge data. The tables that are merged include: 'Vehicle Information', 'Vehicle Lead', 'In-store Maintenance', 'Customer Information', 'Appointment', and 'Lead Resources'.

By utilizing the VIN code and other relevant keys, data from these tables is integrated into a unified dataset, allowing for a comprehensive analysis of vehicle-related information, customer interactions, maintenance records, and appointment details.

2) Data Masking

This project uses real business data, and data masking is required to protect the privacy of client.

We performed the following processing steps:

① Deleted the 'Customer Name' and 'Customer Phone' data that are not related to this analysis.

② Since the VIN code can uniquely identify the vehicle, we masked the VIN code after merging the data.

Table 2-1 VDP Data Masking

E	F
VIN	SourceofLeads
LFV**543	Hua Ao
LFV**442	Hua Ao
LFV**867	Hua Ao
LFV**346	Both Hua Ao and FAW Audi
LFV**269	Hua Ao
LFV**396	Hua Ao
LFV**658	Both Hua Ao and FAW Audi
LFV**543	Hua Ao

3) Data Cleaning

We conducted two aspects of data cleansing. One was to take out the data that did not fit the business logic and the other was to remove duplicate data.

① We found that some vehicles belong to non-private vehicles, or group-owned vehicles, which do not fit our customer invitation logic, so we exclude them.

Table 2-2 VDP Data Cleaning 1

B	C	D
VIN	**Estimated Invitation Data**	**Remark**
LFV**784	2023-05-18	Group-owned
LFV**076	2023-05-18	Group-owned
LFV**0611	2023-05-18	Group-owned
LFV**050	2023-05-18	Group-owned
LFV**181	2023-05-18	Group-owned

② Since the vehicle lead data in the VDP is merged from multiple third-party sources, there is some overlap in vehicle data. For this subset of data, we retain only one record and indicate in the 'Data Source' column that the data is a shared source from both parties.

Table 2-3 VDP Data Cleaning 2

LFV**217	Audi China
WAU**012	Hua Ao
LFV**191	FAW Audi
LFV**255	Both Audi China and Hua Ao
LFV**253	Hua Ao
LFV**214	Both Audi China and FAW Audi

4) Data Imputation

We identified some missing vehicle information data and need to perform data supplementation. Based on the distribution of existing vehicle information data, we will fill in the missing data. Details of how to do this can be found in the attached R program.

5) Data Encoding

We utilized Excel to encode data for the 4 stages of the lead invitation process: 'Connected', 'Invitation', 'Appointment' and ‘In.storeMaintenance’. We used '0' to represent unsuccessful outcomes and '1' to represent successful outcomes.

Table 2-4 VDP Data Encoding

G	H	I
Connected	Intention	Appointment
1	1	1
1	1	1
0	0	0
0	0	0

6) Data Calculation

The vehicle maintenance types we are analyzing include First Maintenance and Regular Maintenance. Due to the differences in maintenance types, the criteria used to identify these two categories of vehicle leads are also distinct. In this project, we calculated the ‘Mileage’ and ‘Vehicle Age’ of Regularly Maintenance vehicles before analysis.

Table 2-5 VDP Data Cleaning 2

Maintenance Type	Mileage	Vehicle Age
First	(Uncalculated)	(Uncalculated)
Regular	‘Mileage’ – ‘Last In-shop Visit Mileage’	‘Purchase Date’ – ‘Last in-store Visit Date’

2.3 VDP Data Overview

This section will show an overview of the vehicle data from the VDP pilot phase.

This pilot campaign lasts for 1 month. Firstly, looking at the maintenance types, regular maintenance takes the majority, with a total of 1,642 leads, while first maintenance has 123 data. This ratio aligns with the basic scenario of vehicle sales and maintenance.

Figure 2-3 Maintenance Type

In terms of geographical distribution, as shown in Figure 2-4, the pilot took place in Beijing, and a large portion of the data during the pilot was collected from Hua Ao, a regional vehicle dealer in Beijing. Consequently, there is a significant regional difference in vehicle data, with most vehicles were in Shunyi and Tongzhou districts in Beijing.

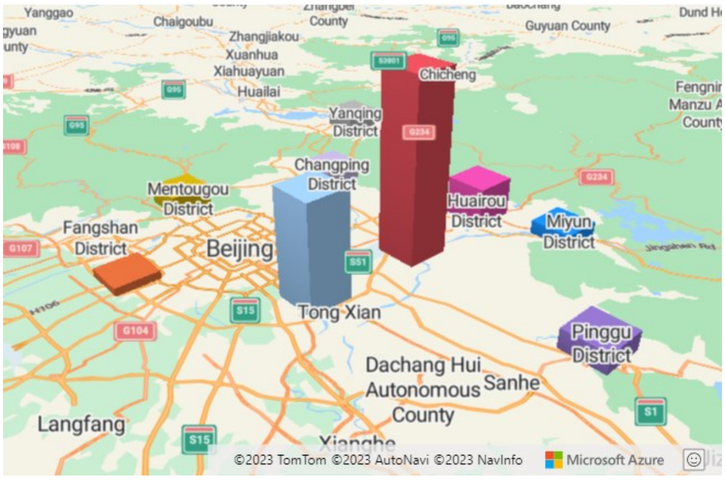

Figure 2-4 Regional Leads Volume

Regarding the conversion of vehicle invitation leads, as shown in Figure 2-5, during this pilot phase there were a total of 1,765 vehicle leads, with 1,420 invitation data. Out of these, 464 customers expressed an intention for maintenance during the invitation process, and 124 customers made maintenance appointments during the pilot. The total in-store maintenance appointments during the pilot period amounted to 250.

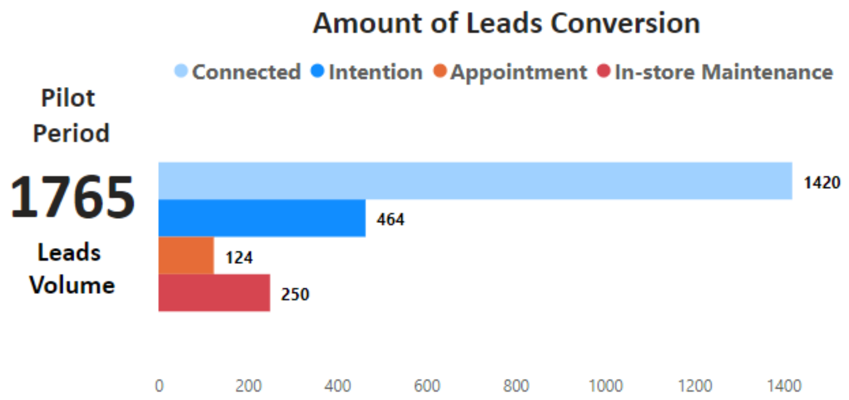

Figure 2-5 Regional Leads Volume

Looking at the distribution of vehicle information, we selected two typical characteristics: mileage and vehicle age. As shown in Figure 2-6 and 2-7, we can see that in this pilot phase, mileage is concentrated in the range of 5,000 to 45,000, while the age of vehicles is mainly within the range of 0.5 to 2 years.

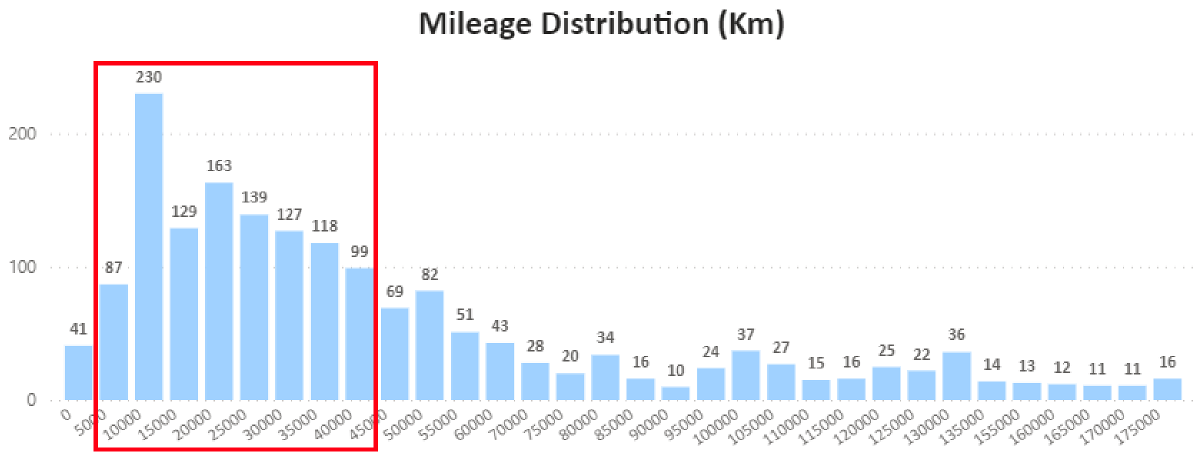

Figure 2-6 VDP Mileage Distribution (Km)

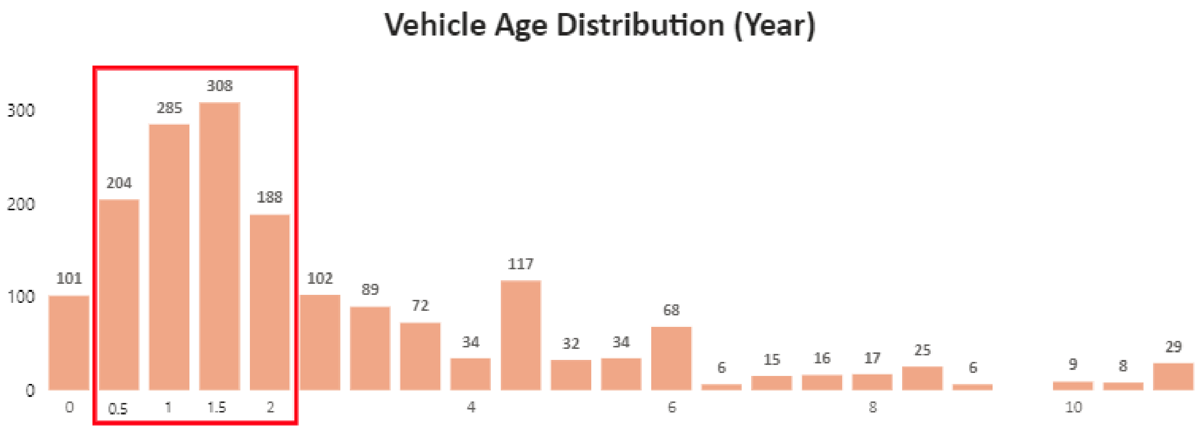

Figure 2-7 VDP Vehicle Age Distribution (Year)

3. Analyzing Processes & Solutions

3.1 Analysis Framework

1) Business Procedure

The major income of Hua Ao comes from after-sale maintenance, and One of the main methods of Hua Ao to remind vehicle owners to do maintenance is through telemarketing. The salesman will each out to the customer on the list. Some customers will pick up the sales call, and the salesman will ask whether they have the intention to visit the dealership and have the maintenance, if the customers say yes, the salesman will book an appointment and the customers could come to the store and take the service. From the procedure mentioned above, we get some data (total lead data, connected data, intention data, appointment data and in-store maintenance data) listed below:

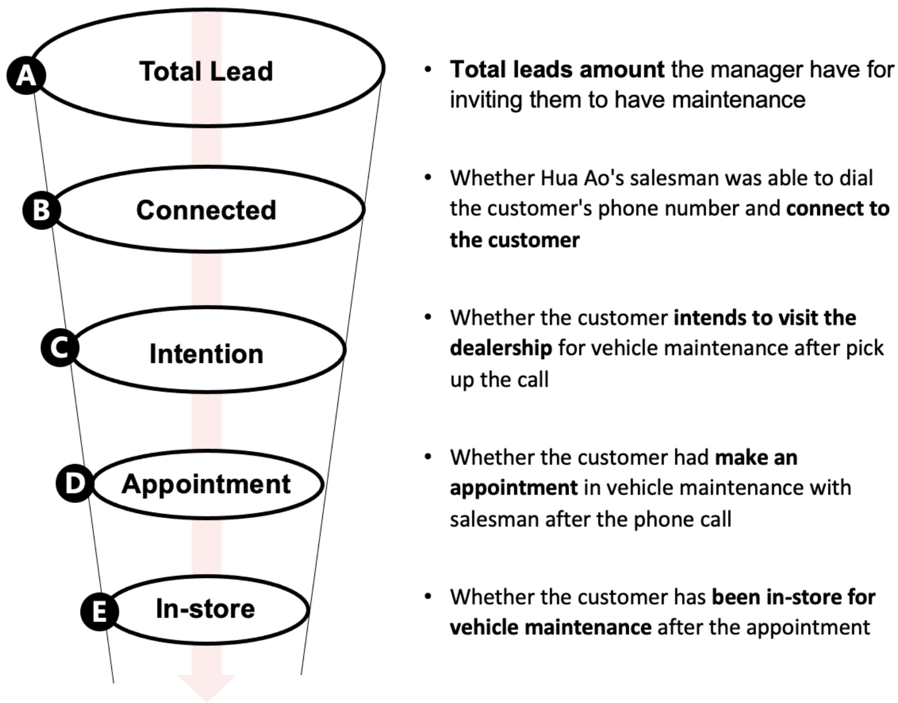

Figure 3-1 Business Procedure

2) Key Matrix

$\text{Connected Rate} = \frac{B}{A}$ $\text{Intention Rate} = \frac{C}{A}$ $\text{Appointment Rate} = \frac{D}{A}$

$\text{In-store Maintenance Rate} = \frac{E'}{A} *$

The connection rate, intention rate, and appointment rate are calculated by the corresponding amount divided by the total lead amount. These first three metrics are used to evaluate the efficiency of the telemarketing. The phone call will have a long-term effect, for example, some customers may have an impression of Hua Ao and come to the store when they are in need, or some customers will gradually become regular customers of Hua Ao. Also, we want to see the overall performance of Hua Ao, so we used the total in-store maintenance divided by the total lead amount to calculate the in-store maintenance rate, but not those come to store after the phone call and appointment.

Finally, we get four metrics: connection rate, intention rate, and appointment rate, and in-store maintenance rate.

3.2 Solution of Insufficient Lead Volume

From Figure 3-2 and Figure 3-3, we can see that, originally, customer data were held separately among three parties. Audi China and FAW Audi both have more than 300 extra data. If gathering all the data and put them into a shared platform (VDP), the available customer data for Hua Ao will increase from 1109 to 1765, which is about a 60% percentage increase.

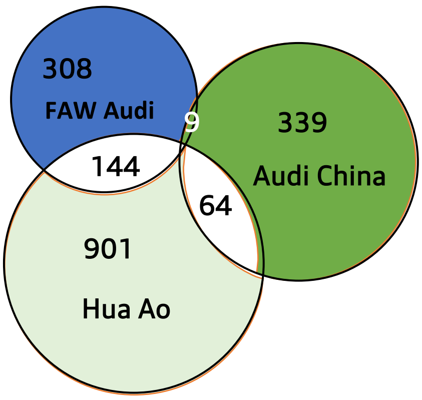

Figure 3-2 Three Parties Leads

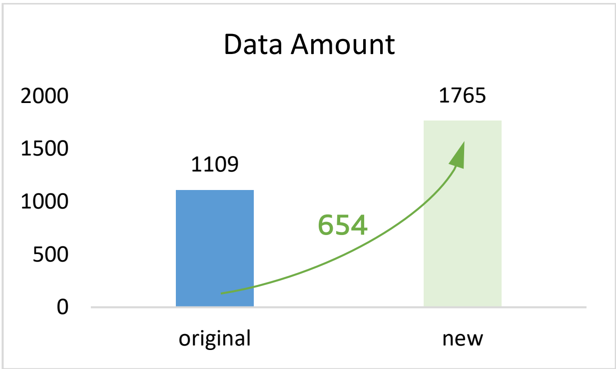

Figure 3-3 Data Amount

What's more, from Figure 3-4 and Figure 3-5, we find that the red bars which represent Audi China usually have the largest value of the 4 metrics I mentioned a few minutes ago, so the data quality of Audi China is the highest. If we calculate the in-store maintenance rate before and after adding in the data from Audi China and FAW Audi, the rate increased from 12% to 17%, that's huge progress.

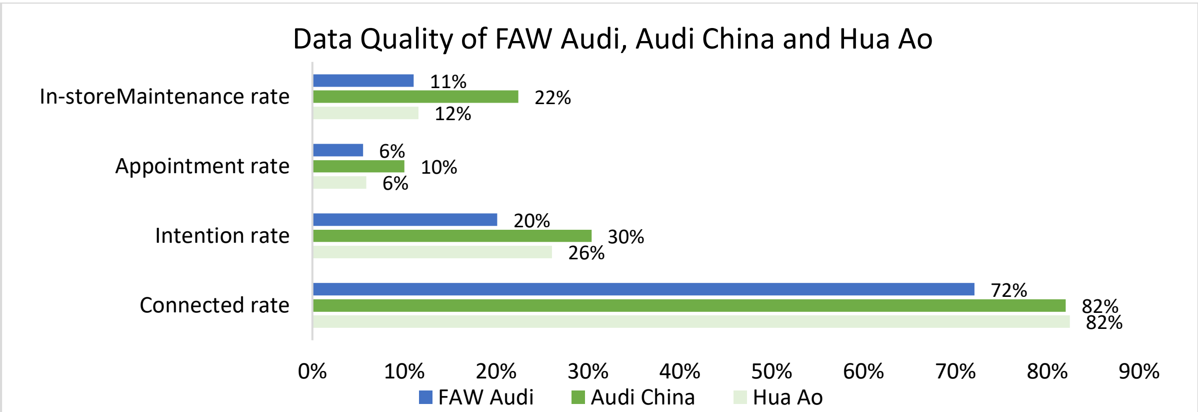

Figure 3-4 Three Parties Data Quality

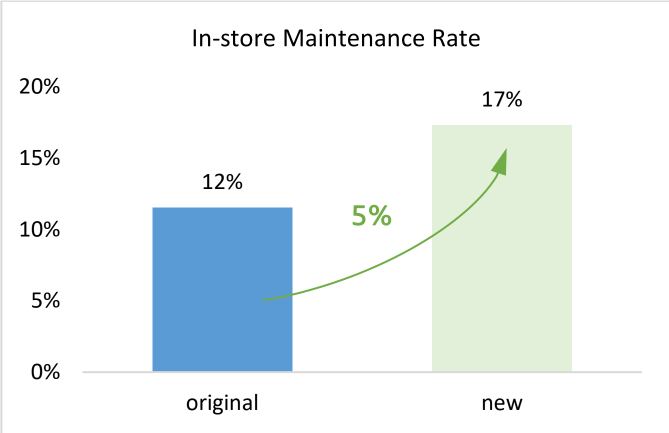

Figure 3-5 In-store Maintenance Rate

3.3 Solution of Inefficiency in Tele-Interviews

In terms of business problem 2, inefficiency in Tele-Interviews, we intend to leverage advanced data modelling techniques to solve this problem, and this process contains two parts.

The first one is using clustering to identify the key factors that influence customers' decisions to visit the dealership for maintenance, and the second is to use decision tree to identify the range within which key factors that affect customer entry have the highest probability of customer entry, which is the threshold for these factors.

1) Clustering to identify the key factors

We chose 8 factors in the clustering process, including 'MaintenanceType', 'Connected', 'Intention,' 'In-storeMaintenance', 'Appointment', 'Salesman', 'Mileage' and 'VehicleAge', but unfortunately, the results were not satisfactory, we set the number of clustering as 4, and among these eight factors, only 'MaintainType' showed a significant correlation with 'In-storeMaintenance.'

We can see that in the last row of clustering, the result of 'MaintenanceType = First' is 0.14, while the other three groups are close to 0. When 'MaintenanceType = Regular', the result is 0.85, and the other three groups are close to 1. However, it is not possible to proceed with the following step based solely on this result.

Table 3-1 Clustering Results

In-storeMaintenance=1	Mileage	VehicleAge	MaintenanceType=First	MaintenanceType=Regular
0.0318258	40829.05	2.01576	5.55112e-17	1
0.0423729	144811.03	2.19737	-2.77556e-17	1
0.0732984	92451.41	2.07984	-2.77556e-17	1
0.246799	11571.49	1.59717	0.14319	0.85681

Therefore, based on business knowledge and experience, we have identified 'Mileage' and 'VehicleAge' as two key factors that can affect customers' decisions to visit the dealership for maintenance, and proceed with the following step.

2) Decision Tree to Find the Threshold

We separated the two maintenance types. In the first case, which is Regular Maintenance, we divided the training set and test set, adjusted parameters, and constructed a decision tree.

For regular maintenance leads, Figure 3-6 showed under the conditions of 'Vehicle Age' greater than 0.21, less than or equal to 1.29, and 'Mileage' less than or equal to 9408, customers have a 78.3% probability of entering the store for maintenance.

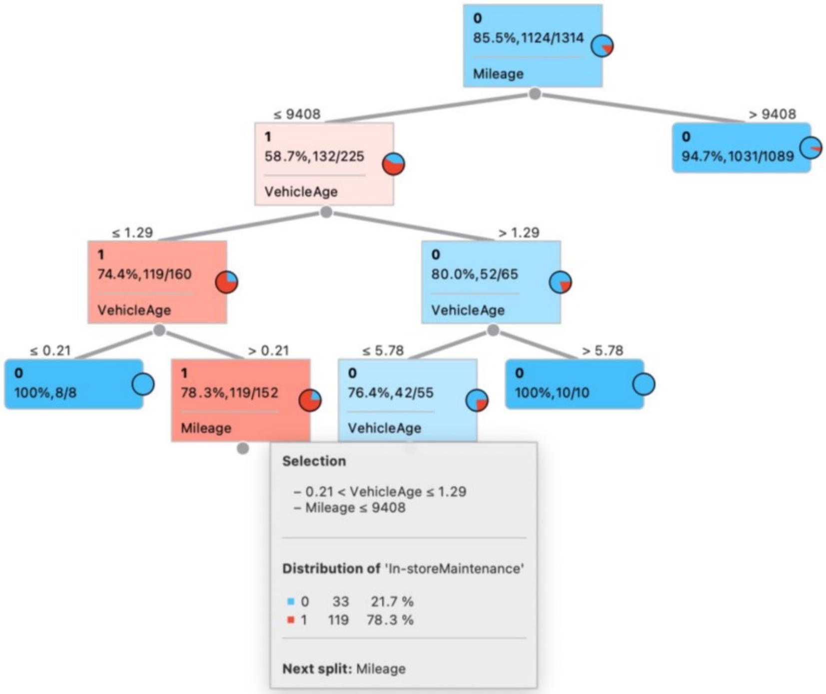

Figure 3-6 Regular Maintenance Decision Tree

For first maintenance leads, Figure 3-7 showed under the conditions of 'Vehicle Age' greater than 0.27, when the 'Mileage' is greater than 7942, customers have a 60% probability of entering the store for maintenance. However, in this case, there are only 5 available data in the dataset.

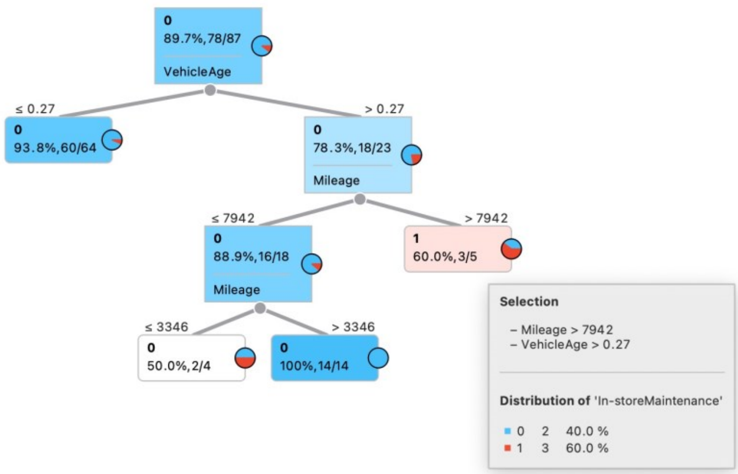

Figure 3-7 First Maintenance Decision Tree

However, due to the small amount of data, the analysis results with the First Maintenance are not reliable. Therefore, we will mainly focus on the situation where the Maintenance Type is Regular.

In the dataset with regular maintenance, the probability of customers entering the store for maintenance was 14.1%, which is a very low probability. However, after obtaining the new rule through decision tree modelling (As shown in Table 3-2), we applied the new rule to the original dataset, selected customers who are more likely to enter the store for maintenance, and we focus on these customers, then we obtained a new probability of customers entering the store for maintenance of 78.3%, which is significantly higher than before.

Table 3-2 Comparison of Customer Entry Rate Before and After

Maintenance = Regular	Before	After
In-store rate	14.1%	78.3%

This indicates that our modelling is very effective and greatly improves the customer entering rate, At the same time, it greatly improves the work efficiency of dealer staff.

4. Summary & Prospects

4.1 Summary

For question 1, we addressed the issue of insufficient lead quantity for Hua Ao by combining the leads from Hua Ao, Audi China, and FAW Audi.

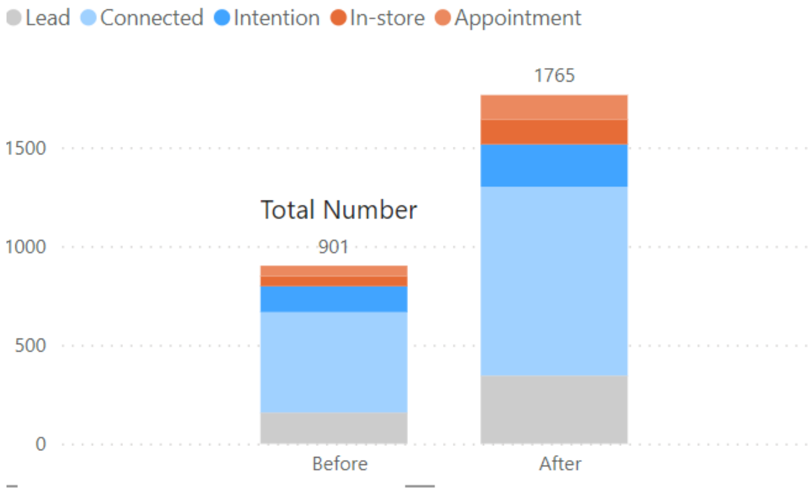

Figure 4-1 Change in Lead Quantity after Sharing Data

As shown in the bar chart, after consolidating data from these three platforms, we observed significant increases in in-store, appointment, intention, connection, and overall lead numbers. The lead count surged from 901 to 1765.

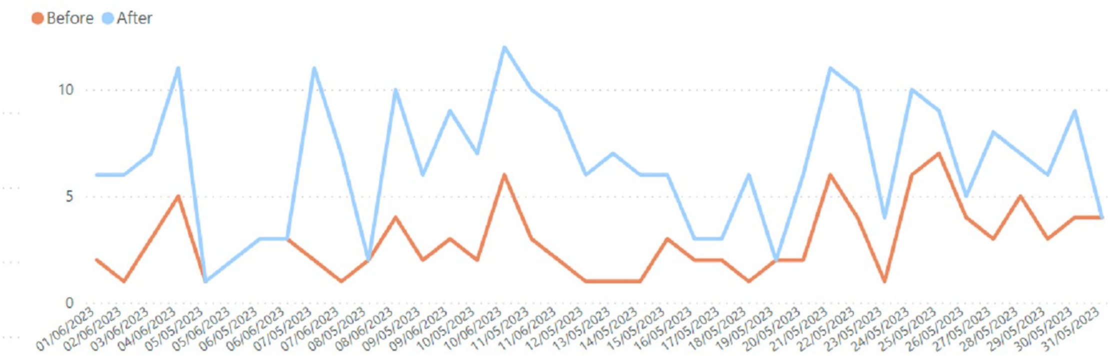

Figure 4-2 Change in Lead Quantity with Time

Moreover, the line chart illustrates that the blue line consistently remains above the orange line, indicating a continuous daily increase of in-store quantity after sharing data.

For Question 2, we employed a decision tree to establish filtering rules for both regular and first maintenance data to identify the target clients with whom we intend to contact.

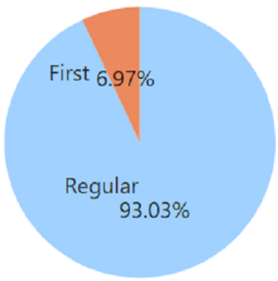

Figure 4-3 Maintenance Type Proportion

Our analysis revealed that the first maintenance data constituted 6.97% percentage of the whole data, while the regular maintenance data accounted for 93.03% percent.

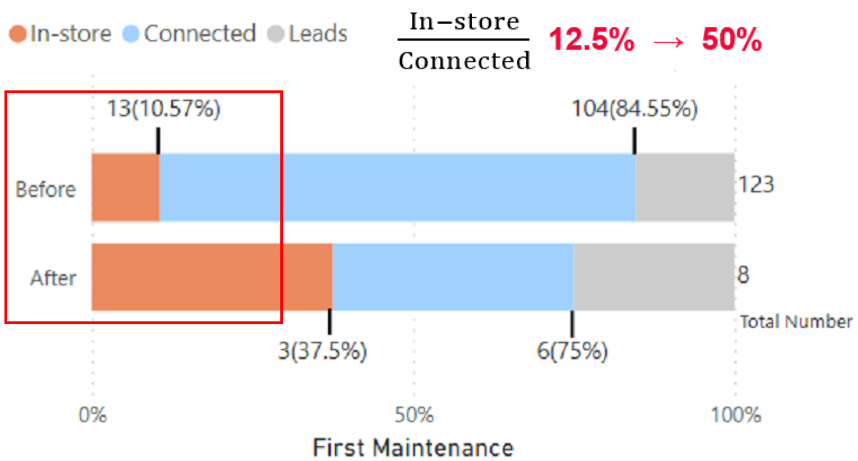

Figure 4-4 Change in Lead Conversion Efficiency for First Maintenance

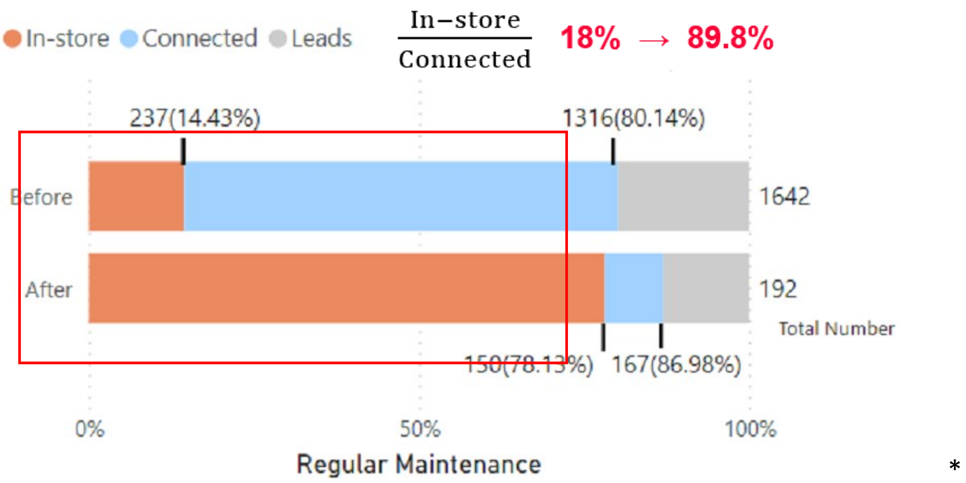

Figure 4-5 Change in Lead Conversion Efficiency for Regular Maintenance

Following the application of our filtering rules, we observed a significant enhancement in lead conversion efficiency, which means the ratio of the number of entries to the number of connections. We can see the stacked chart on the left side that for first maintenance, the entry rate increase from 10.57% to 37.5%. While the lead conversion efficiency increased from 12.5% to 50%. And from the stacked chart one the right side, we can see the entry rate increased from 14.43% to 78.13%. And the lead conversion efficiency increased from 18% to 89.8%.

Overall, after building VDP, we increase the lead volume as well as the lead conversion efficiency.

4.2 Project Timeline

4.3 Prospects of VDP

First, we will continue to improve the model. Since we only built the model with one-month data, we will continuously collect data in subsequent operations and adjust model.

Second, we will explore more fields that VDP can be used, such as the Analysis of Solicitation of Alarm Clues, and Future Potential Opportunities and Revenue.

5. Reference Materials

[1] http://www.caam.org.cn/chn/4/cate_38/con_5236151.html

[2] McKinsey Greater China - Dealers Transformation

[3] https://cn.gtadata.com/