

Spring 2022 Database Project

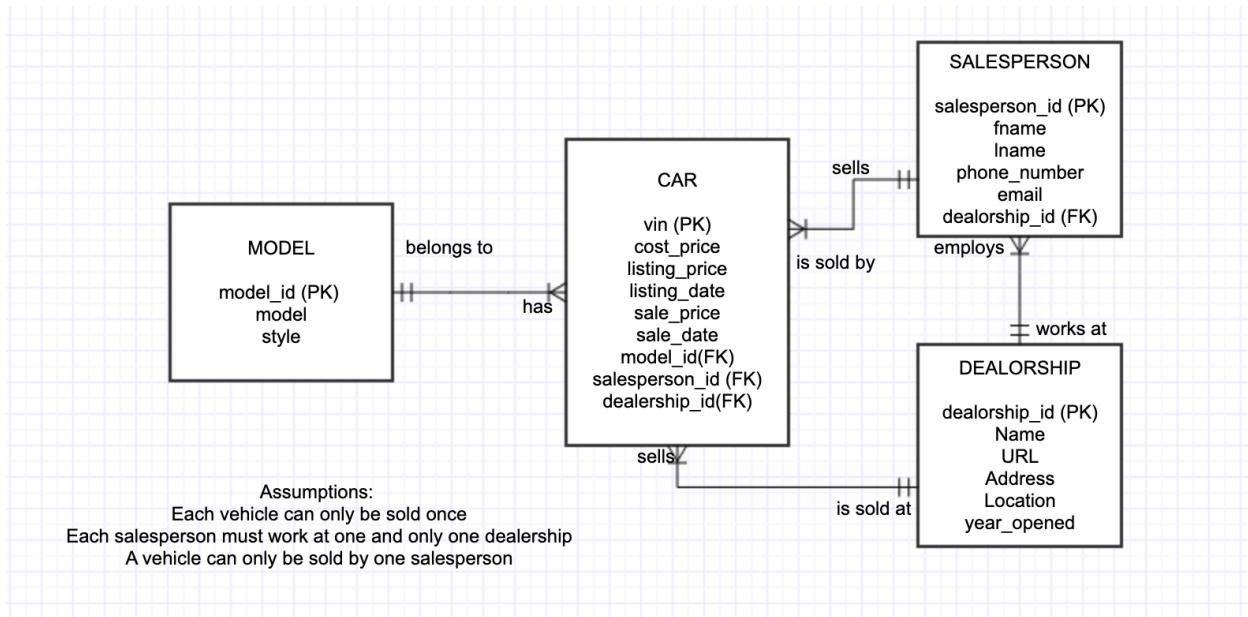
Cadillac Sales Data

By Tianze Shou



Part A: Data Modeling, Creation and Basic Queries

Entity Relationship Diagram:



Data Dictionary:

CAR			
Field	Data Type	Description	Example Data
vin (PK)	VARCHAR (200)	Vehicle Identification Number	1GKLVKED8AJ155580
cost_price	INT	Base price received from Cadillac	30,000
listing_price	INT	Asking price	40,000
listing_date	DATE	Listing date	2020-12-12
sale_price	INT	Actual price the vehicle was sold	35,000
sale_date	DATE	Date where the transaction was made	2020-12-30
salesperson_id (FK)	INT	Unique ID for salesperson related to this transaction	1352
dealership_id (FK)	INT	Unique ID for dealership associated with the sale of the vehicle	1234

SALESPERSON			
Field	Data Type	Description	Example Data
salesperson_id (PK)	INT	Unique ID for every salesperson	1352
fname	VARCHAR (200)	First name of salesperson	Jack
lname	VARCHAR (200)	Last name of salesperson	Daniels
phone	VARCHAR (200)	Phone number of the salesperson	4129667599
email	VARCHAR (200)	Email address of the salesperson	123@321.com
dealership_id (FK)	INT	Unique ID of the dealership associated with the salesperson	1234

DEALERSHIP			
Field	Data Type	Description	Example Data
dealership_id (PK)	INT	Unique ID for each Cadillac dealership	1234
name	VARCHAR (200)	Name of the dealership	North Bay Cadillac
url	VARCHAR (200)	Website URL	https://www.northbaycadillac.com/
address	VARCHAR (200)	Address of the dealership	720 Northern Blvd, Great Neck, NY 11021
location	VARCHAR (200)	Location of the dealership	Great Neck
year_opened	INT	Year opened for the dealership	1980

MODEL			
Field	Data Type	Description	Example Data
model_id (PK)	INT	Unique ID for each model	1234
model	VARCHAR (200)	Model name	CT4
style	VARCHAR (200)	Body style	Sedan

DEALERSHIP:

#	Name	Type	Collation	Attributes	Null	Default	Extra	Action
<input type="checkbox"/>	1	dealership_id	int(11)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	2	name	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	3	url	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	4	address	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	5	location	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	6	year	year(4)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More

SELECT * FROM `DEALERSHIP`

☐ Profiling
 [\[Edit inline \]](#)
[\[Edit \]](#)
[\[Explain SQL \]](#)
[\[Create PHP code \]](#)
[\[Refresh \]](#)

☐ Show all Number of rows: 25 Filter rows: Search this table

Sort by key: None

+ Options

<div><div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>		dealership_id	name	url	address	location	year
<div><div><div></div><div></div><div></div></div></div> <div><div><div>Edit</div><div>Copy</div><div>Delete</div></div></div>		2000	North Bay Cadillac	https://cadillac-northbay.com	720 Northern Blvd, Great Neck, NY 11021	Great Neck	1980
<div><div><div></div><div></div><div></div></div></div> <div><div><div>Edit</div><div>Copy</div><div>Delete</div></div></div>		2001	Grand Prize Cadillac	https://cadillac-grandprize.com	20 NY-304, Nanuet, NY 10954	Nanuet	1978
<div><div><div></div><div></div><div></div></div></div> <div><div><div>Edit</div><div>Copy</div><div>Delete</div></div></div>		2002	Mount Kisco Cadillac	https://cadillac-mountkisco.com	175 N Bedford Rd Suite B, Mt Kisco, NY 10549	Mt Kisco	1991
<div><div><div></div><div></div><div></div></div></div> <div><div><div>Edit</div><div>Copy</div><div>Delete</div></div></div>		2003	Robert Basil Cadillac	http://www.cadillacrobertbasil.com	3475 Southwestern Blvd, Orchard Park, NY 14127	Orchard Park	1992
<div><div><div></div><div></div><div></div></div></div> <div><div><div>Edit</div><div>Copy</div><div>Delete</div></div></div>		2004	Bical Cadillac	http://www.cadillac-bical.com	2859 Flatbush Ave, Brooklyn, NY 11234	Brooklyn	1985
<div><div><div></div><div></div><div></div></div></div> <div><div><div>Edit</div><div>Copy</div><div>Delete</div></div></div>		2005	Paul Conte Cadillac	http://www.cadillac-paulconte.com	169 W Sunrise Hwy, Freeport, NY 11520	Freeport	2001

MODEL:

#	Name	Type	Collation	Attributes	Null	Default	Extra	Action
<input type="checkbox"/>	1	model_id	int(11)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	2	model	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	3	style	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More

`SELECT * FROM `MODEL``

☐ Profiling [\[Edit inline \]](#) [\[Edit \]](#) [\[Explain SQL \]](#) [\[Create PHP code \]](#) [\[Refresh \]](#)

☐ Show all | Number of rows: | Filter rows:

Sort by key:

+ Options

		model_id	model	style
<input type="checkbox"/>	Edit Copy Delete	1400	CT4	sedan
<input type="checkbox"/>	Edit Copy Delete	1450	CT4-V	sedan
<input type="checkbox"/>	Edit Copy Delete	1500	CT5	sedan
<input type="checkbox"/>	Edit Copy Delete	1550	CT5-V	sedan
<input type="checkbox"/>	Edit Copy Delete	2400	XT4	suv
<input type="checkbox"/>	Edit Copy Delete	2500	XT5	suv
<input type="checkbox"/>	Edit Copy Delete	2600	XT6	suv

SALESPERSON:

#	Name	Type	Collation	Attributes	Null	Default	Extra	Action
<input type="checkbox"/>	1	salesperson_id	int(11)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	2	fname	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	3	lname	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	4	phone	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	5	email	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	6	dealership_id	int(11)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More

`SELECT * FROM `SALESPERSON``

☐ Profiling [\[Edit inline \]](#) [\[Edit \]](#) [\[Explain SQL \]](#) [\[Create PHP code \]](#) [\[Refresh \]](#)

☐ Show all | Number of rows: | Filter rows:

Sort by key:

+ Options

		salesperson_id	fname	lname	phone	email	dealership_id
<input type="checkbox"/>	Edit Copy Delete	101	Moses	Crafford	(212) 555-6183	mcrafford@cadillac.com	2001
<input type="checkbox"/>	Edit Copy Delete	102	Jack	Clobburn	(212) 555-9348	jclobburn@cadillac.com	2004
<input type="checkbox"/>	Edit Copy Delete	103	Gerbard	Forrester	(212) 555-9912	gforrester@cadillac.com	2001
<input type="checkbox"/>	Edit Copy Delete	104	Sophia	Hobburn	(212) 555-3246	shobburn@cadillac.com	2002
<input type="checkbox"/>	Edit Copy Delete	105	Edward	Josh	(212) 555-9448	ejosh@cadillac.com	2000
<input type="checkbox"/>	Edit Copy Delete	106	Gabriel	King	(212) 555-2455	gking@cadillac.com	2005
<input type="checkbox"/>	Edit Copy Delete	107	Travis	King	(212) 555-2456	tking@cadillac.com	2000
<input checked="" type="checkbox"/>	Edit Copy Delete	108	Silver	Santiago	(212) 555-3385	ssantiago@cadillac.com	2003

VEHICLE:

#	Name	Type	Collation	Attributes	Null	Default	Extra	Action
<input type="checkbox"/>	1	vin	varchar(200)	utf8_bin	No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	2	list_price	decimal(10,0)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	3	sale_price	decimal(10,0)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	4	list_date	date		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	5	sale_date	date		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	6	cost_price	decimal(10,0)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	7	model_id	int(11)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	8	salesperson_id	int(11)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More
<input type="checkbox"/>	9	dealership_id	int(11)		No	None		Change Drop Primary Unique Index Spatial Fulltext Distinct values More

SELECT * FROM `VEHICLE`

☐ Profiling [\[Edit inline \]](#) [\[Edit \]](#) [\[Explain SQL \]](#) [\[Create PHP code \]](#) [\[Refresh \]](#)

1

> >>

Number of rows: 25

Filter rows:

Sort by key:

None

+ Options

		vin	list_price	sale_price	list_date	sale_date	cost_price	model_id	salesperson_id	dealership_id
<input type="checkbox"/>	Edit Copy Delete	CT4GE11RXBLA13262	37066	36242	2021-03-01	2021-03-01	32947	1400	109	2002
<input type="checkbox"/>	Edit Copy Delete	CT4GE11RXBLA13265	40366	39469	2021-01-01	2021-01-05	35881	1400	103	2001
<input type="checkbox"/>	Edit Copy Delete	CT4GE11RXBLA13268	34010	33254	2021-06-01	2021-06-01	30231	1400	106	2005
<input type="checkbox"/>	Edit Copy Delete	CT4GE11RXBLA13274	34201	33441	2021-01-01	2021-01-01	30401	1400	112	2005
<input type="checkbox"/>	Edit Copy Delete	CT4GE11RXBLA13275	40412	39514	2021-01-01	2021-01-01	35921	1400	112	2005
<input type="checkbox"/>	Edit Copy Delete	CT4GE11RXBLA13276	33985	33230	2021-01-01	2021-01-14	30209	1400	112	2005
<input type="checkbox"/>	Edit Copy Delete	CT4GE11RXBLA13279	42097	41162	2021-01-01	2021-01-01	37420	1400	105	2000

Part B: Queries

What was the average profit or loss (sale price – cost price) over all vehicles sold (over all the dealerships in the dataset)?

```
SELECT ROUND(AVG(sale_price - cost_price),2) AS 'Average Profit/Loss Overall' FROM VEHICLE
```

☐ Profiling [Edit inline] [Edit] [Explain SQL] [Create PHP code] [Refresh]

1 | Number of rows: 25 | Filter rows: Search this table

Sort by key: None

+ Options

Average Profit/Loss Overall

4858.23

What was the average markup (list price – cost price) of all vehicles sold (over all the dealerships in the dataset)?

```
SELECT ROUND(AVG(list_price - cost_price ),2) AS 'Average Markup Overall' FROM VEHICLE
```

☐ Profiling [Edit inline] [Edit] [Explain SQL] [Create PHP code] [Refresh]

1 | Number of rows: 25 | Filter rows: Search this table

Sort by key: None

+ Options

Average Markup Overall

6072.79

What was the average markup (list price – cost price) for each model of the vehicles sold (over all the dealerships in the dataset)?

```
SELECT ROUND(AVG(VEHICLE.list_price - VEHICLE.cost_price),2) AS 'Average Markup', MODEL.model AS 'Model' FROM VEHICLE INNER JOIN MODEL ON VEHICLE.model_id = MODEL.model_id GROUP BY MODEL.model ORDER BY AVG(VEHICLE.list_price - VEHICLE.cost_price)
```

☐ Profiling [Edit inline] [Edit] [Explain SQL] [Create PHP code] [Refresh]

☐ Show all | Number of rows: 25 | Filter rows: Search this table

Sort by key: None

+ Options

Average Markup	Model
3963.77	CT5
4157.72	CT4
5088.45	XT4
5498.83	XT5
5719.36	CT4-V
6243.76	XT6
6597.26	CT5-V
9684.19	Escalade

What was the average time on market for each model (over all the dealerships in the dataset)?

```
SELECT AVG(datediff(VEHICLE.sale_date,VEHICLE.list_date)) AS 'Average Time on Market (days)', MODEL.model AS 'Model' FROM VEHICLE INNER JOIN MODEL ON VEHICLE.model_id = MODEL.model_id GROUP BY MODEL.model ORDER BY AVG(datediff(VEHICLE.sale_date,VEHICLE.list_date))
```

☐ Profiling

[\[Edit inline \]](#)

[\[Edit \]](#)

[\[Explain SQL \]](#)

[\[Create PHP code \]](#)

[\[Refresh \]](#)

☐ Show all

Number of rows: 25

Filter rows:

+ Options

Average Time on Market (days)	Model
6.0533	Escalade
6.5600	XT4
6.9250	XT6
6.9300	CT4
6.9800	CT5-V
7.5233	CT4-V
9.5850	CT5
10.1700	XT5

What was the total number of vehicles sold for each model for each dealership?

```
SELECT COUNT(VEHICLE.vin) AS 'Number of Vehicles Sold', MODEL.model AS 'Model', DEALERSHIP.name AS 'Dealership' FROM VEHICLE INNER JOIN MODEL ON VEHICLE.model_id = MODEL.model_id INNER JOIN DEALERSHIP ON VEHICLE.dealership_id = DEALERSHIP.dealership_id GROUP BY MODEL.model, DEALERSHIP.name ORDER BY VEHICLE.model_id, COUNT(VEHICLE.vin)
```

☐ Profiling

[\[Edit inline \]](#)

[\[Edit \]](#)

[\[Explain SQL \]](#)

[\[Create PHP code \]](#)

[\[Refresh \]](#)

1

>

>>

☐ Show all

Number of rows: 25

Filter rows:

+ Options

Number of Vehicles Sold	Model	Dealership
15	CT4	Bical Cadillac
20	CT4	Robert Basil Cadillac
21	CT4	North Bay Cadillac
37	CT4	Mount Kisco Cadillac
43	CT4	Grand Prize Cadillac
64	CT4	Paul Conte Cadillac
37	CT4-V	Bical Cadillac
40	CT4-V	Robert Basil Cadillac
53	CT4-V	Paul Conte Cadillac
55	CT4-V	Grand Prize Cadillac

Which salesperson made the average highest profit or loss (sale price – cost price)?

```
SELECT ROUND(AVG(VEHICLE.sale_price - VEHICLE.cost_price), 2) AS 'Average Profit/Loss', SALESPERSON.salesperson_id AS 'ID',  
SALESPERSON.fname AS 'First Name', SALESPERSON.lname AS 'Last Name' FROM VEHICLE INNER JOIN SALESPERSON ON VEHICLE.salesperson_id =  
SALESPERSON.salesperson_id GROUP BY SALESPERSON.salesperson_id ORDER BY AVG(VEHICLE.sale_price - VEHICLE.cost_price) DESC
```

☐ Profiling [\[Edit inline \]](#) [\[Edit \]](#) [\[Explain SQL \]](#) [\[Create PHP code \]](#) [\[Refresh \]](#)

☐ Show all | Number of rows: 25 | Filter rows:

Sort by key:

None

+ Options

Average Profit/Loss	ID	First Name	Last Name
5057.28	111	Vogel	Valeria
5029.62	110	Liz	Santiago
5001.86	105	Edward	Josh
4892.69	112	Radcliffe	Wesley
4882.97	102	Jack	Clobburn
4854.00	100	Chadwick	Michael

Custom query: what are the numbers of vehicles sold and average profit/loss for each body style (sedan, suv) over all vehicles and all dealerships?

Showing rows 0 – 1 (2 total, Query took 0.0050 seconds.)

```
SELECT COUNT(VEHICLE.vin) AS "Number of Vehicles Sold", ROUND(AVG(VEHICLE.sale_price - VEHICLE.cost_price), 2) AS "Average Profit", MODEL.style AS "Body  
Style" FROM VEHICLE INNER JOIN MODEL ON VEHICLE.model_id = MODEL.model_id GROUP BY MODEL.style
```

☐ Profiling [\[Edit inline \]](#) [\[Edit \]](#) [\[Explain SQL \]](#) [\[Create PHP code \]](#) [\[Refresh \]](#)

☐ Show all | Number of rows: 25 | Filter rows:

+ Options

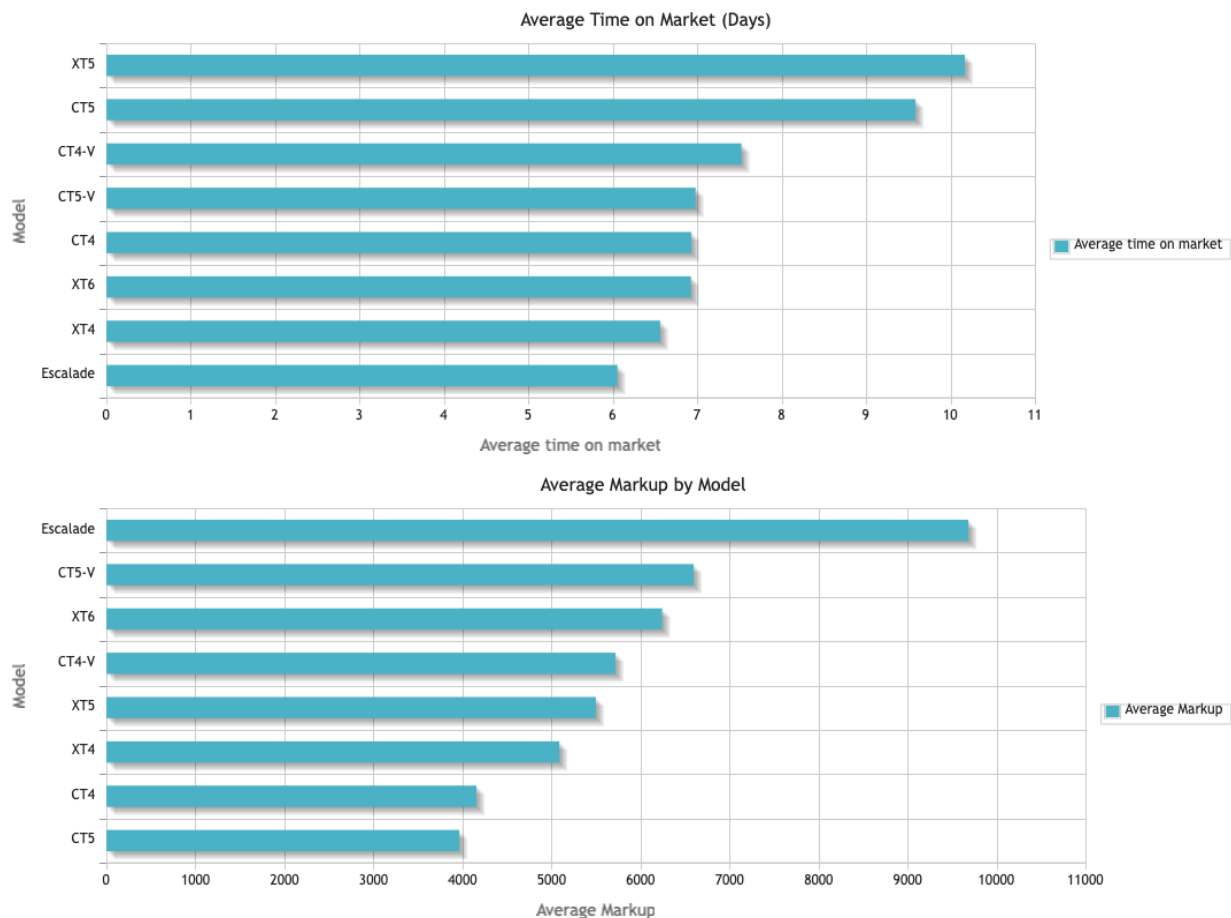
Number of Vehicles Sold	Average Profit	Body Style
900	4141.85	sedan
900	5574.60	suv

Part C: Business Memo

To: Cadillac U.S.A.
From: Tianze Shou
Date: April 16, 2022
Subject: Model Promotion and Dealership Expansion

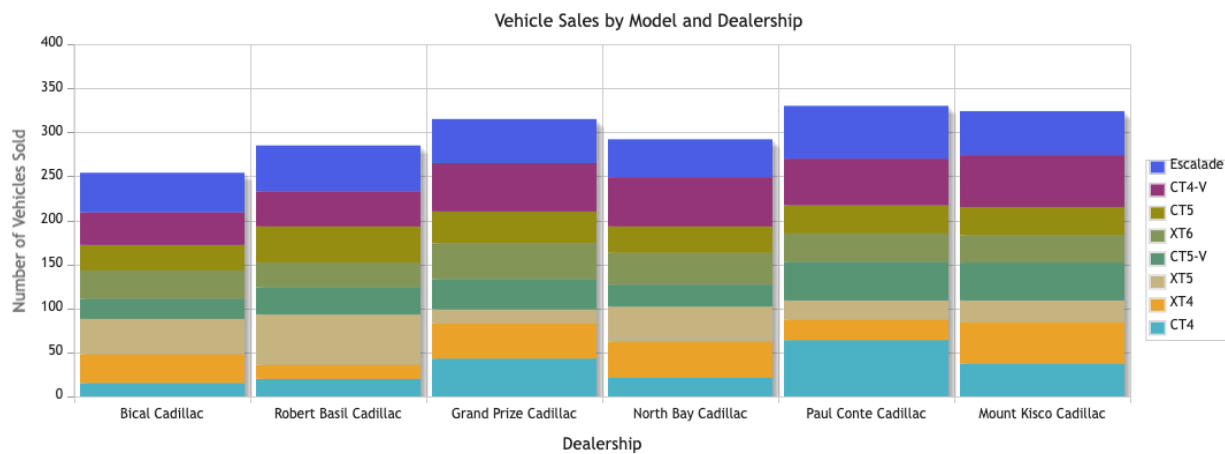
Dear Cadillac U.S.A. Executives,

After thorough examination on Cadillac sales data from various dealerships in the greater New York region, I would recommend enhancing the production level of the Escalade model.



As one can observe from the first graph above, the Escalade model stays in the market for the shortest amount of time. This signifies that its demand is higher than the other model while production is lower. The second graph shows that the markup of Escalade far exceeds the others, so the dealerships can profit the most from Escalade than any other models. Therefore, based on the strong demand and the profitability of this model, the production of Escalade should be promoted.

After some analysis on the sales data of each dealership, the conclusion is that the one with the most urgent need to expand is Paul Conte Cadillac.



Not only does Paul Conte retain the largest number of total vehicle sales, this dealership also has the largest number of Escalade sales. If the corporation is going to promote the production of Escalade models in the future, Paul Conte Cadillac could take the leadership in selling these promoted products. Therefore, Paul Conte Cadillac should be expanded.

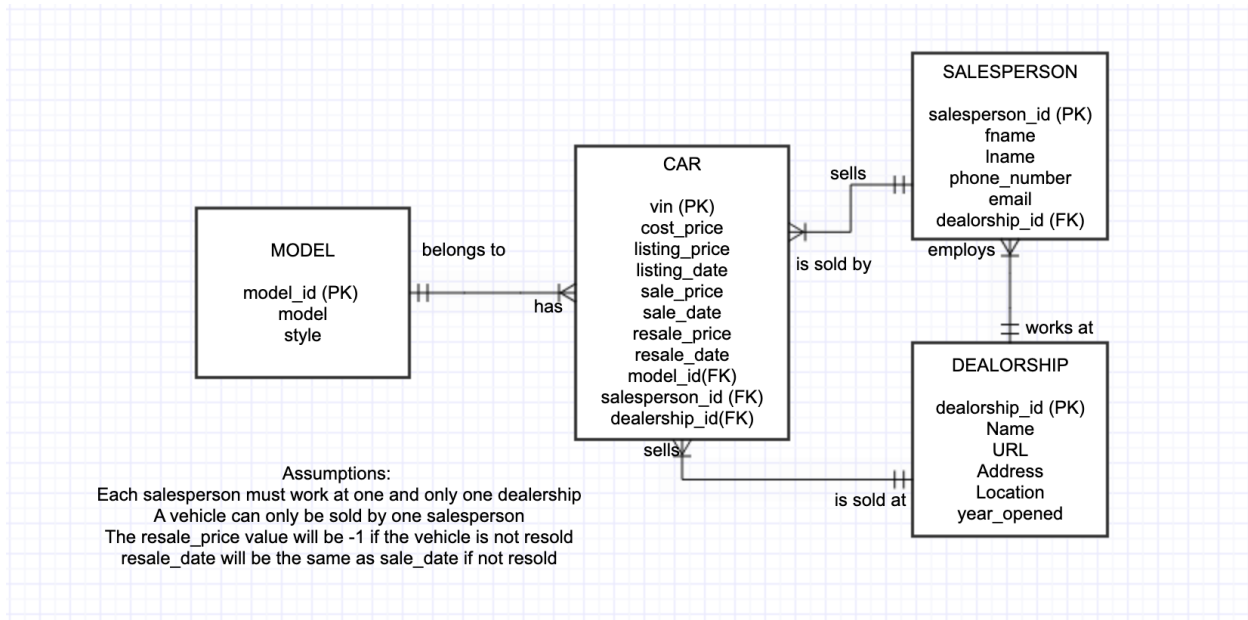
In short, the corporation should consider promoting the sales of the Escalade model and expanding Paul Conte dealership in the New York region.

Regards,

Tianze Shou
Cadillac U.S.A. IS Analyst

Part D: Additional Data Considerations

Modified ERD:



Example queries:

1. The savings involved in purchasing pre-owned vehicles can be assessed by querying the average difference between sale_price and resale_price for each model over all dealerships.
2. The returning rate of each salesperson.
3. The average time in days between sale and resale for each model for all dealerships.

Example SQL query (1):

```
SELECT ROUND(AVG(VEHICLE.sale_price - VEHICLE.resale_price), 2) AS 'Saving',  
        MODEL.model AS 'Model'  
FROM VEHICLE  
INNER JOIN MODEL  
ON VEHICLE.model_id = MODEL.model_id  
GROUP BY MODEL.model  
ORDER BY AVG(VEHICLE.sale_price - VEHICLE.resale_price)  
WHERE VEHICLE.resale_price > 0
```