

Smart Factories with SAP Project Presentation

Car Sales Analysis in America Using SAP Analytics Cloud



Aziz Fahmi (30461097)
Umberto Corvino (30018392)

Agenda

1. Introduction
2. Services and Tools
3. Data Integration
4. Python
5. SAC
6. Next steps

Introduction

This project analyzes car sales data in America using SAP Analytics Cloud (SAC).

We will explore:

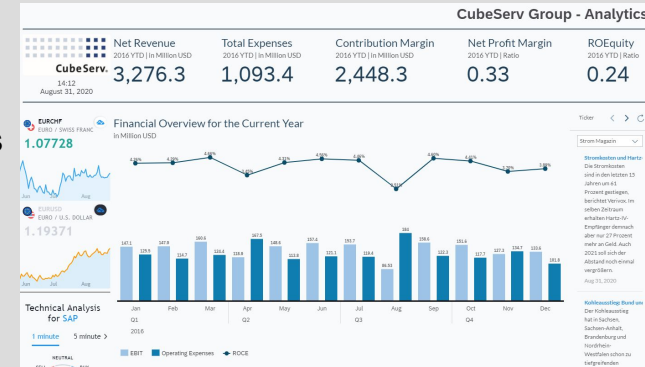
- trends in car prices
- manufacturer performance
- other key factors influencing car sales.



Services and tools

SAP Analytics Cloud (SAC)

- ❑ **Primary data analysis platform:** Analyses data and creating interactive dashboards and reports
- ❑ **Visualization:** Transforms complex data into easy-to-understand graphics and charts
- ❑ **Building analytic applications:** Develops custom apps to meet specific business needs and enhance decision-making



Python

- ❑ **Initial data cleaning:** Ensures raw data is free of errors and inconsistencies
- ❑ **Data preprocessing:** Performs steps like normalization and transformation to prepare data for analysis



Data Integration

Data Source:

- **Kaggle dataset:** Utilizes the "Car Sales Data in America" dataset from Kaggle



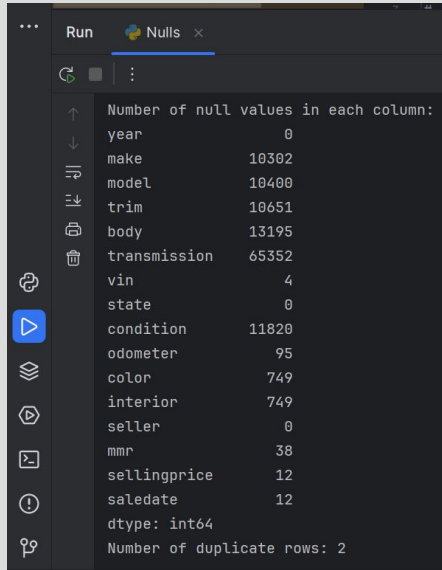
Data Storage and import:

- **Cleaned data in CSV format then imported in SAC:** for creating insightful reports and dashboards



Python

Raw data



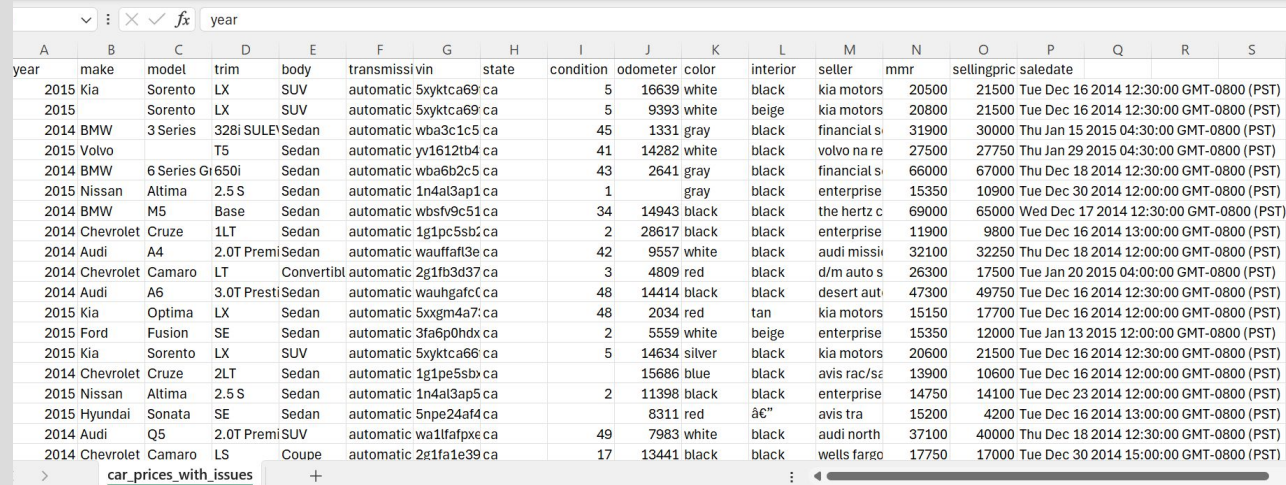
Run Nulls

Number of null values in each column:

year	0
make	10302
model	10400
trim	10651
body	13195
transmission	65352
vin	4
state	0
condition	11820
odometer	95
color	749
interior	749
seller	0
mmr	38
sellingprice	12
saledate	12
dtype: int64	

Number of duplicate rows: 2

Null values and duplicates



year

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
year	make	model	trim	body	transmission	vin	state	condition	odometer	color	interior	seller	mmr	sellingprice	saledate			
2015	Kia	Sorento	LX	SUV	automatic	5xyktca69 ca		5	16639	white	black	kia motors	20500	21500	Tue Dec 16 2014 12:30:00 GMT-0800 (PST)			
2015	Sorento	LX	SUV	automatic	5xyktca69 ca			5	9393	white	beige	kia motors	20800	21500	Tue Dec 16 2014 12:30:00 GMT-0800 (PST)			
2014	BMW	3 Series	328i SULE	Sedan	automatic	wba3c1c5 ca		45	1331	gray	black	financial s	31900	30000	Thu Jan 15 2015 04:30:00 GMT-0800 (PST)			
2015	Volvo		T5	Sedan	automatic	yv1612tb4 ca		41	14282	white	black	volvo na re	27500	27750	Thu Jan 29 2015 04:30:00 GMT-0800 (PST)			
2014	BMW	6 Series G	650i	Sedan	automatic	wba6b2c5 ca		43	2641	gray	black	financial s	66000	67000	Thu Dec 18 2014 12:30:00 GMT-0800 (PST)			
2015	Nissan	Altima	2.5 S	Sedan	automatic	1n4al3ap1 ca		1		gray	black	enterprise	15350	10900	Tue Dec 30 2014 12:00:00 GMT-0800 (PST)			
2014	BMW	M5	Base	Sedan	automatic	wbstfv9c51 ca		34	14943	black	black	the hertz c	69000	65000	Wed Dec 17 2014 12:30:00 GMT-0800 (PST)			
2014	Chevrolet	Cruze	1LT	Sedan	automatic	1g1pc5sb7 ca		2	28617	black	black	enterprise	11900	9800	Tue Dec 16 2014 13:00:00 GMT-0800 (PST)			
2014	Audi	A4	2.0T Premi	Sedan	automatic	wauffaf13e ca		42	9557	white	black	audi missi	32100	32250	Thu Dec 18 2014 12:00:00 GMT-0800 (PST)			
2014	Chevrolet	Camaro	LT	Convertibl	automatic	2g1fb3d37 ca		3	4809	red	black	d/m auto s	26300	17500	Tue Jan 20 2015 04:00:00 GMT-0800 (PST)			
2014	Audi	A6	3.0T Presti	Sedan	automatic	wauhgaf1c ca		48	14414	black	black	desert aut	47300	49750	Tue Dec 16 2014 12:30:00 GMT-0800 (PST)			
2015	Kia	Optima	LX	Sedan	automatic	5xxgm4a7 ca		48	2034	red	tan	kia motors	15150	17700	Tue Dec 16 2014 12:00:00 GMT-0800 (PST)			
2015	Ford	Fusion	SE	Sedan	automatic	3fa6p0hdx ca		2	5559	white	beige	enterprise	15350	12000	Tue Jan 13 2015 12:00:00 GMT-0800 (PST)			
2015	Kia	Sorento	LX	SUV	automatic	5xyktca66 ca		5	14634	silver	black	kia motors	20600	21500	Tue Dec 16 2014 12:30:00 GMT-0800 (PST)			
2014	Chevrolet	Cruze	2LT	Sedan	automatic	1g1pe5sbv ca			15686	blue	black	avis rac/se	13900	10600	Tue Dec 16 2014 12:00:00 GMT-0800 (PST)			
2015	Nissan	Altima	2.5 S	Sedan	automatic	1n4al3ap5 ca		2	11398	black	black	enterprise	14750	14100	Tue Dec 23 2014 12:00:00 GMT-0800 (PST)			
2015	Hyundai	Sonata	SE	Sedan	automatic	5npae24af4 ca			8311	red	âc	avis tra	15200	4200	Tue Dec 16 2014 13:00:00 GMT-0800 (PST)			
2014	Audi	Q5	2.0T Premi	SUV	automatic	wa1fafpxe ca		49	7983	white	black	audi north	37100	40000	Thu Dec 18 2014 12:30:00 GMT-0800 (PST)			
2014	Chevrolet	Camaro	LS	Coupe	automatic	2s1fa1e39 ca		17	13441	black	black	wells fargo	17750	17000	Tue Dec 30 2014 15:00:00 GMT-0800 (PST)			

car_prices_with_issues

558837 rows

16 columns

Abbreviations are displayed in state column

Wrong date format

Python

Results after Python code

```
Data cleaning complete. Cleaned data saved
Number of null values in each column:
year          0
make          0
model         0
trim          0
body          0
transmission  0
state         0
condition     0
odometer      0
color         0
interior      0
seller        0
mmr           0
sellingprice  0
saledate      0
dtype: int64
Number of duplicate rows: 0
```

No more **Null** values and **duplicates**

P135071	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
135071	2006	Toyota	Sienna	XLE Limite	Minivan	automatic	5tdza22c4	Georgia	25	109540	white	beige	nalley toy	7975	7300	01/22/2015	
135072	2006	Volkswage	Jetta	TDI	Sedan	automatic	3wvdt71kx	Georgia	34	126901	gray	gray	santander	4525	5800	01/20/2015	
135073	2005	BMW	5 Series	525i	Sedan	automatic	wbana535	Maryland	22	97922	gray	gray	sheehy for	7000	6300	01/27/2015	
135074	2005	Chevrolet	Colorado	Z85 LS Bas	Crew Cab	automatic	1gcds1367	California	35	160982	black	gray	honda of p	5600	6000	01/20/2015	
135075	2005	Chrysler	Town and (Touring	Minivan	automatic	2c4gp54l8	Georgia		28	160614	silver	gray	nalley toy	2150	1600	02/03/2015	
135076	2005	Chrysler	Sebring	Base	Convertibl	automatic	1c3e145x6	Georgia	2	89472	silver	gray	nalley toy	2000	1500	01/20/2015	
135077	2005	Ford	F-150	XLT	SuperCrew	automatic	1ftpw1259	California	1	108119	silver	gray	tradewind	8300	7900	01/20/2015	
135078	2005	Ford	Escape	XLT	SUV	automatic	1fmyu9317	Pennsylvania	3	115645	blue	gray	faulkner su	4125	4200	01/19/2015	
135079	2005	Ford	Excursion	Limited	SUV	automatic	1fmsu43p	California	36	186262	beige	beige	hemborg fr	10150	9700	01/20/2015	
135080	2005	Ford	F-150	XLT	SuperCab	automatic	1ftrx12w5	Texas	3	160049	black	tan	java enterj	5275	3100	01/20/2015	
135081	2005	Ford	F-150	XLT	SuperCrew	automatic	1ftpw1453	California	27	197067	gray	gray	hemborg fr	4600	6300	01/20/2015	
135082	2005	Ford	Freestyle	SE	Wagon	automatic	1fmzk0112	Georgia	27	161843	beige	beige	santander	1950	1400	01/20/2015	
135083	2014	Chrysler	Town and (Touring	Minivan	automatic	2c4rc1bg0	Maryland		2	25967	beige	brown	enterprise	18700	17100	01/20/2015	
135084	2005	Honda	Odyssey	Touring	Minivan	automatic	5fnrl3889	Maryland	29	110534	red	tan	tappahanr	7025	6300	02/03/2015	
135085	2005	Jeep	Wrangler	X	SUV	manual	1j4fa39s8	California	26	67892	silver	black	walters au	11150	11600	01/20/2015	
135086	2005	Infiniti	QX56	Base	SUV	automatic	5n3aa08c	Missouri	21	168223	silver	gray	arizona au	7850	6500	01/20/2015	
135087	2005	MINI	Cooper	Base	Convertibl	automatic	wmwrf335	California	39	126850	blue	âc"	e/e auto w	3525	3600	01/20/2015	
135088	2005	Mercedes-	E-Class	E320	Sedan	automatic	wdbuf65j2	California	38	78529	white	gray	walters au	7775	9400	01/20/2015	
135089	2005	Mazda	Tribute	s	SUV	automatic	4f2y041x	Ohio	19	108294	blue	tan	bobb auto	3975	3600	02/10/2015	

From **558837** rows to **472323** rows

Sales date column modified

State column modified (Example from **ca** to **California**)

Importing Data into SAP Analytics Cloud (SAC)

- ❑ **Process:** The cleaned CSV file is imported into SAC, where it becomes accessible for further analysis and visualization
- ❑ **Dataset Overview:** The dataset displays some rows and 15 columns, including fields like year, condition, and odometer

The screenshot displays the SAP Analytics Cloud (SAC) interface. On the left is a navigation menu with options like Home, Files, Apps, Stories, Analytic Applications, Data Analyzer, Digital Boardroom, Datasets, More..., Security, and System. The main area shows the 'clean_car_sales_data' dataset overview. A 'Sampled Data' pop-up window is visible, indicating a full data size of 472322 rows and 15 columns, with a sampled size of 2000 rows. The dataset overview panel on the right shows 'clean_car_sales_data' with 2000 rows and 15 columns. Below this, there are sections for 'Measures (5)' and 'Dimensions (10)'. The measures section includes 'year', 'condition', and 'odometer', each with a 'SUM' aggregation. The dimensions section is partially visible. At the bottom right, there is a 'Validate Full Dataset' button.

Year	Brand	Model	Condition	Odometer
2015	Kia	Niro	Good	10000
2014	BMW	5 Series	Good	15000
2014	BMW	3 Series	Good	12000
2014	BMW	1 Series	Good	8000
2014	Chevrolet	Cruze	Good	10000
2014	Audi	A4	Good	12000
2014	Chevrolet	Camaro	Good	15000
2014	Audi	A6	Good	18000
2015	Kia	Optima	Good	10000
2015	Ford	Fusion	Good	12000
2015	Kia	Sorento	Good	15000

Defining Measures and Dimensions in SAC

- **Measures:** Key metrics like **MMR** (Market Value) and **Selling Price**
- **Dimensions:** Various dimensions such as **year**, **make**, **model**, **trim**, **body**, **transmission**, **state**, **condition**, **odometer**, **color**, **interior**, **seller**, and **saledate**

New features are available. Check them out.

SAP Datasets clean_car_sales_data

File Display Actions

	AA Country/R...	AA trim	AA model	AA make	z2 year	AA body	AA transmissi...	AA vin
1	USA	S	Focus	Ford	2013	Sedan	automatic	1fadp3a29d
2	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f21d
3	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f24d
4	USA	SE	Fiesta	Ford	2013	Sedan	automatic	3fadp4b7d
5	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f23d
6	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f26d
7	USA	Titanium	Focus	Ford	2013	Sedan	automatic	1fadp3f29d
8	USA	Titanium	Fusion	Ford	2013	Sedan	automatic	3fadp0d99c
9	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f25d
10	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f22d
11	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f23d
12	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f29d
13	USA	SE	Fusion	Ford	2013	Sedan	automatic	3fadp0h71d
14	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f21d
15	USA	SE	Fusion	Ford	2013	Sedan	automatic	3fadp0h71d
16	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f22d
17	USA	SE	Focus	Ford	2013	Sedan	automatic	1fadp3f22d

Dataset Overview

clean_car_sales_data 2000 rows 20 columns

Search

Output Columns

Measures (2)

- mmr SUM
- sellingprice SUM

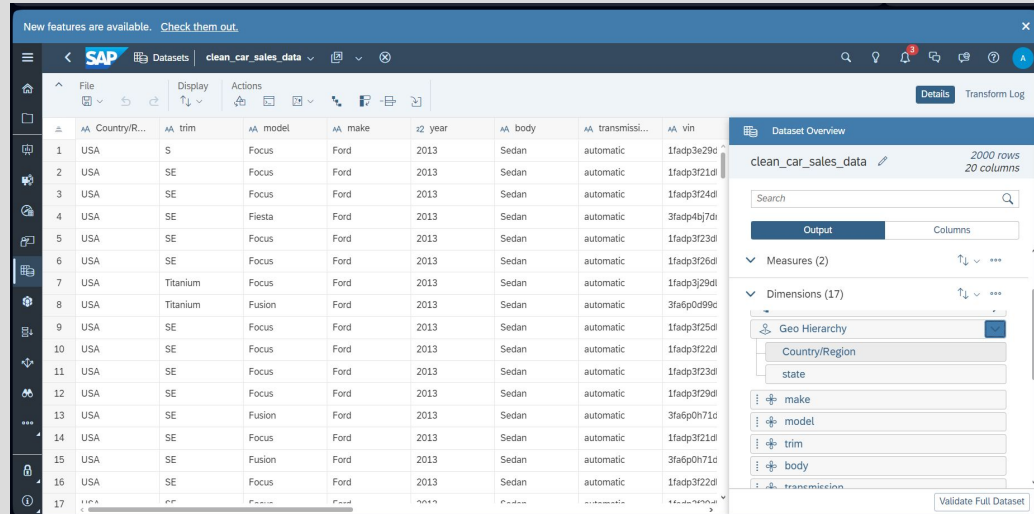
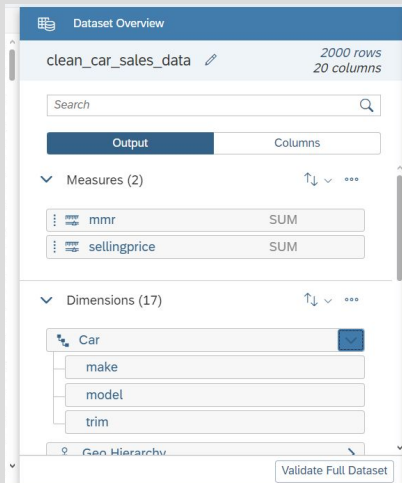
Dimensions (17)

- Car
- Geo Hierarchy
- make
- model
- trim

Valid Show desktop

Hierarchy and Geo Hierarchy Creation

- ❑ **Car Hierarchy:** Hierarchy for car data including **make**, **model**, and **trim**
- ❑ **Geo Hierarchy:** Created to visualize data geographically by state. A new **Country/Region** column is added with the value "USA" and a hierarchical structure is established (Country level 1, State level 2)



Data Analyzer Area

- ❑ **Selection:** Users can select measures and dimensions to be calculated and visualized
- ❑ **Example Analysis:** Ford cars sold for \$128,725,920 in 2015.

New features are available. [Check them out.](#)

SAP Data Analyzer Insight 1 of car sales

clean_car_sales_data

make	Year	state	sellingprice
Ford	2015	Florida	128,725,920
		Texas	121,707,783
		Illinois	93,063,207
		California	91,125,055
		Tennessee	90,032,977
		Missouri	85,806,781
		Michigan	79,717,025
		Georgia	68,649,293
		Pennsylvania	50,225,080
		New Jersey	44,924,934
		Ohio	41,872,511
		Minnesota	33,828,085
		New York	30,901,351
		North Carolina	21,561,453
		Colorado	20,427,155

Available Objects

Search

Measures

fx+ Add Calculation

Select All

mmr

sellingprice

Dimensions

Version

body

color

condition

interior

make

model

odometer

saledate

Builder

Available Objects

Display Type:

Table

Rows

make

Year

state

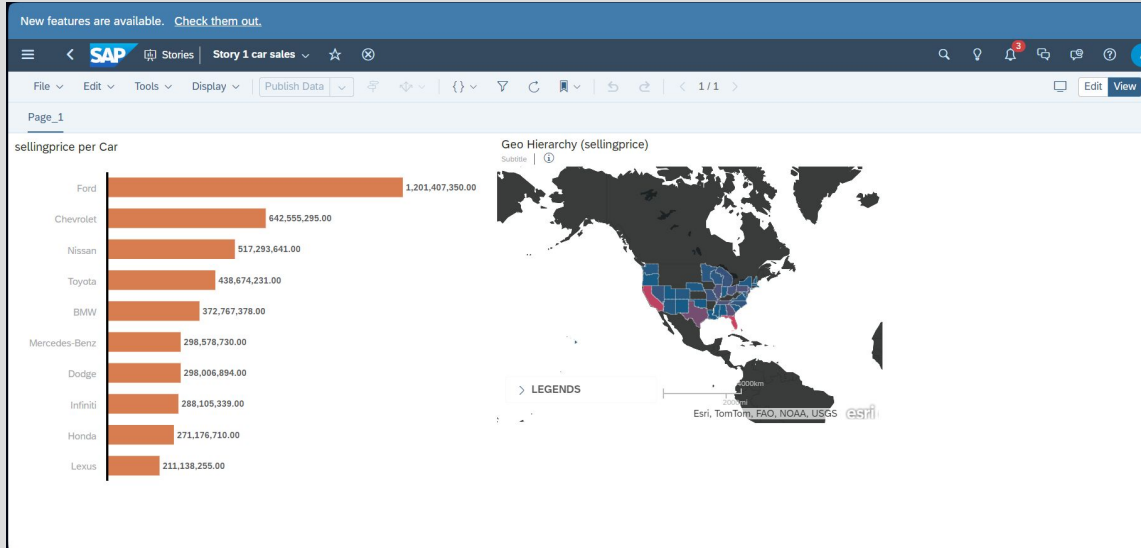
Columns

Measures

sellingprice

Creating Visualizations in SAC

- Visualization Types: Various types of visualizations can be created, including **bar charts**, **pie charts**, and **geo maps**
- Example Visualization: A **bar chart** showing selling prices per car and a **geo map** depicting the sales distribution across states



Next Steps

- ❑ **More Custom Dashboards:** Create more dashboards tailored to user roles each with relevant KPIs and visualizations
- ❑ **More Drill-Down:** Enable users to click on high-level metrics and explore detailed data, such as **sales by dealership or car model**
- ❑ **More Filters and Selectors:** Add interactive filters to customize data views by **date range, region, or car model**

Any questions?
