



# Automobile sales

BUSINESS REPORT 2018 TO 2020

# Data

#	Column	Non-Null Count	Dtype
0	ORDERNUMBER	2747 non-null	int64
1	QUANTITYORDERED	2747 non-null	int64
2	PRICEEACH	2747 non-null	float64
3	ORDERLINENUMBER	2747 non-null	int64
4	SALES	2747 non-null	float64
5	ORDERDATE	2747 non-null	datetime64[ns]
6	DAYS_SINCE_LASTORDER	2747 non-null	int64
7	STATUS	2747 non-null	object
8	PRODUCTLINE	2747 non-null	object
9	MSRP	2747 non-null	int64
10	PRODUCTCODE	2747 non-null	object
11	CUSTOMERNAME	2747 non-null	object
12	PHONE	2747 non-null	object
13	ADDRESSLINE1	2747 non-null	object
14	CITY	2747 non-null	object
15	POSTALCODE	2747 non-null	object
16	COUNTRY	2747 non-null	object
17	CONTACTLASTNAME	2747 non-null	object
18	CONTACTFIRSTNAME	2747 non-null	object
19	DEALSIZE	2747 non-null	object

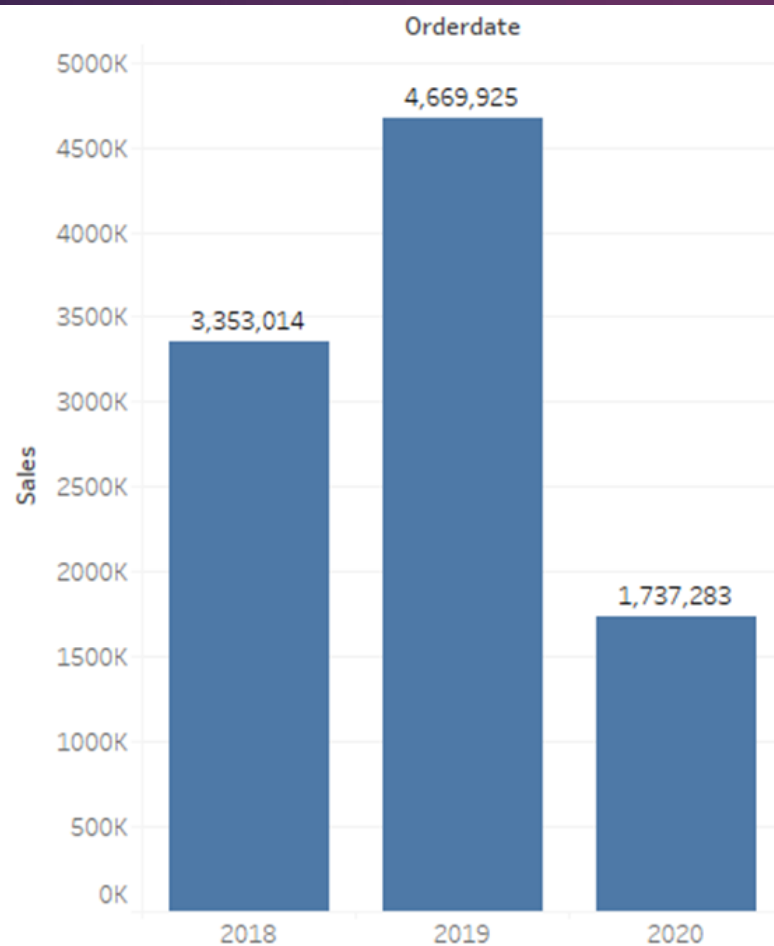
dtypes: datetime64[ns](1), float64(2), int64(5), object(12)  
memory usage: 429.3+ KB

- ▶ The data set has 2747 rows and 20 columns in it
- ▶ The dataset contains 7 integer data type, 1 date column and 12 string data type in it.
- ▶ The data does not have any null values and duplicate values.
- ▶ So the data has been processed using KNIME tool for RFM analysis.

# Data- descriptive stats

	count	mean	std	min	25%	50%	75%	max
<b>ORDERNUMBER</b>	2747.0	10259.761558	91.877521	10100.00	10181.000	10264.00	10334.500	10425.00
<b>QUANTITYORDERED</b>	2747.0	35.103021	9.762135	6.00	27.000	35.00	43.000	97.00
<b>PRICEEACH</b>	2747.0	101.098951	42.042548	26.88	68.745	95.55	127.100	252.87
<b>ORDERLINENUMBER</b>	2747.0	6.491081	4.230544	1.00	3.000	6.00	9.000	18.00
<b>SALES</b>	2747.0	3553.047583	1838.953901	482.13	2204.350	3184.80	4503.095	14082.80
<b>DAYS_SINCE_LASTORDER</b>	2747.0	1757.085912	819.280576	42.00	1077.000	1761.00	2436.500	3562.00
<b>MSRP</b>	2747.0	100.691664	40.114802	33.00	68.000	99.00	124.000	214.00

# Year-on Year Sales

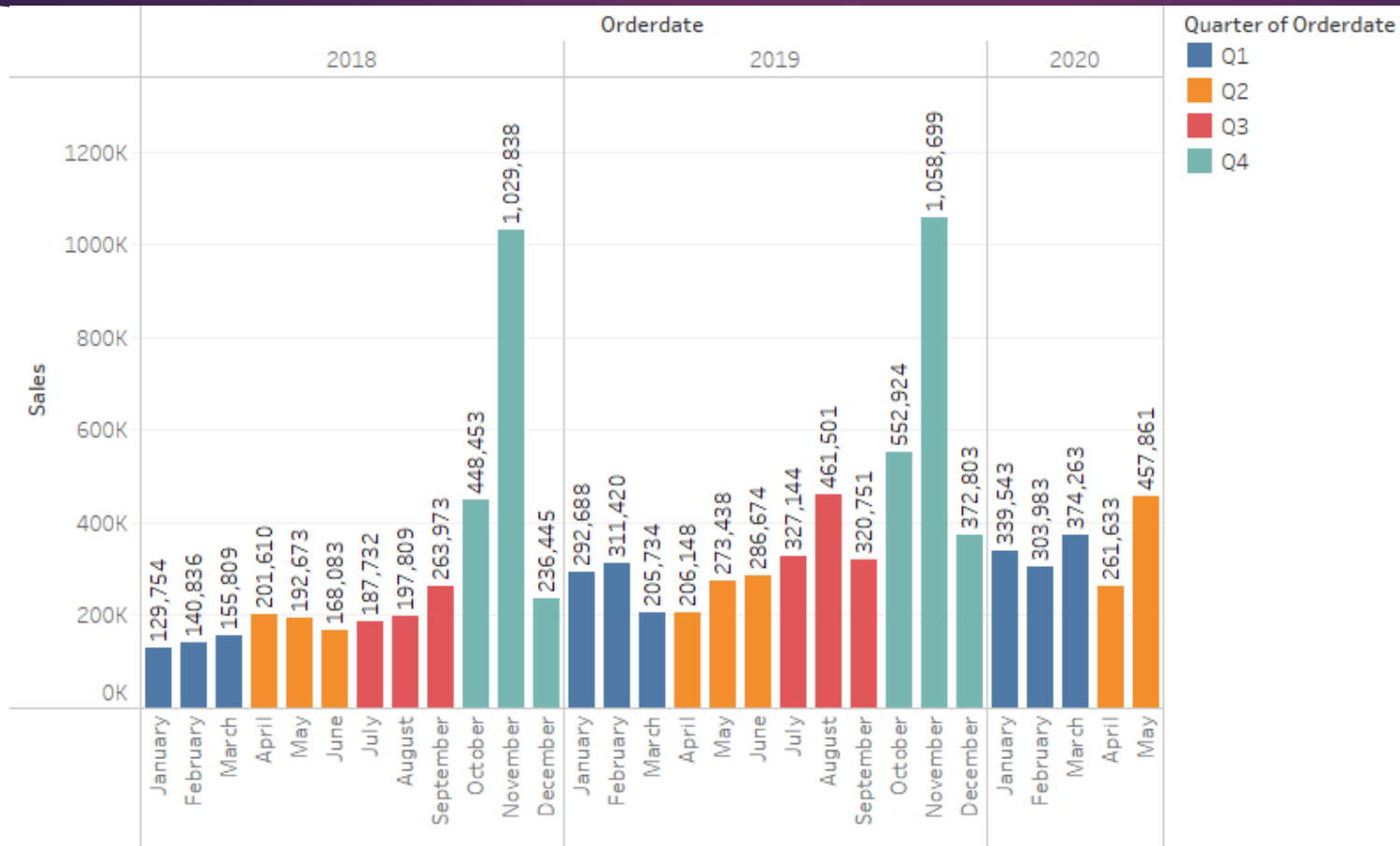


Sum of Sales for each Orderdate Year.

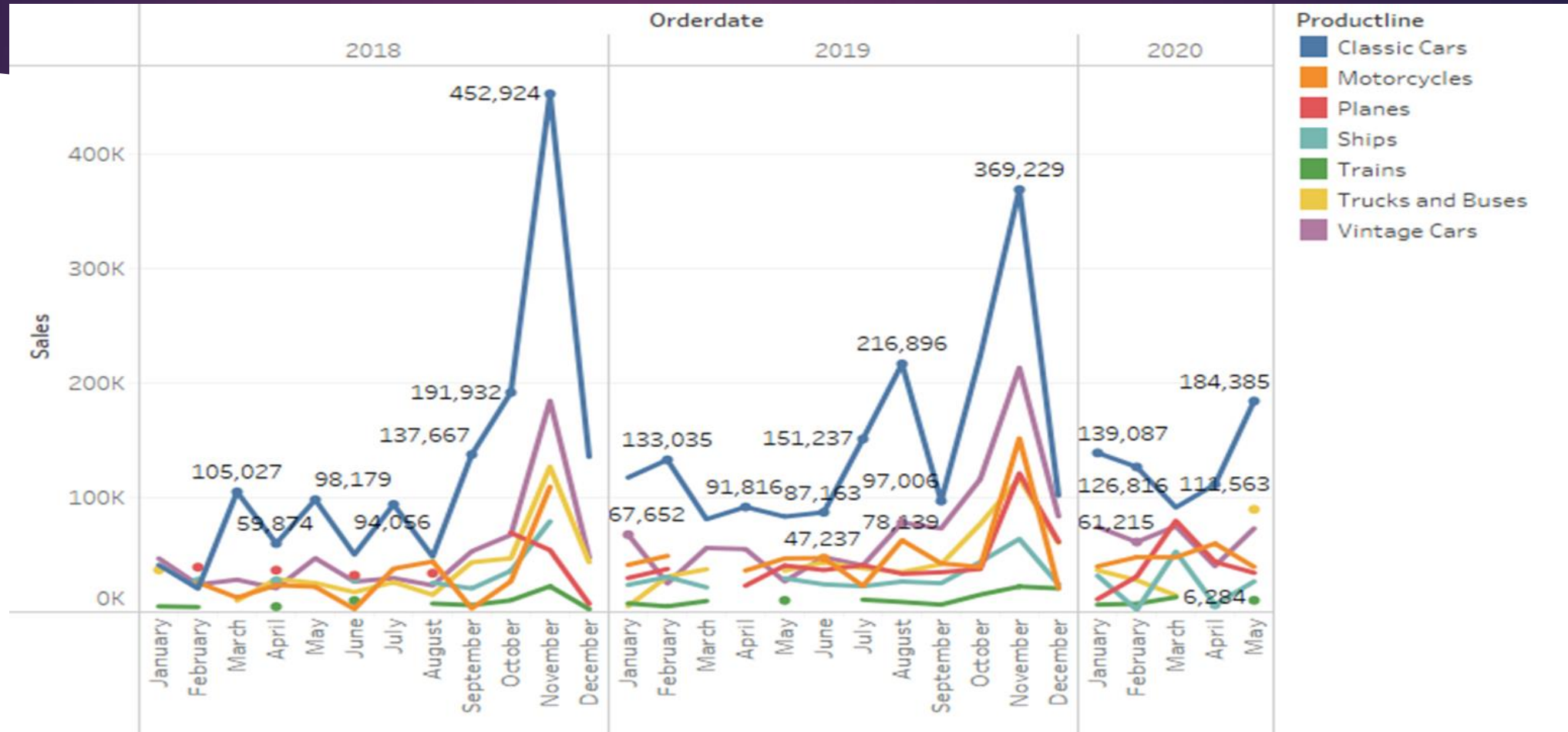


The trend of sum of Sales for Orderdate Month. Color shows details about Orderdate Year.

# Monthly Sales

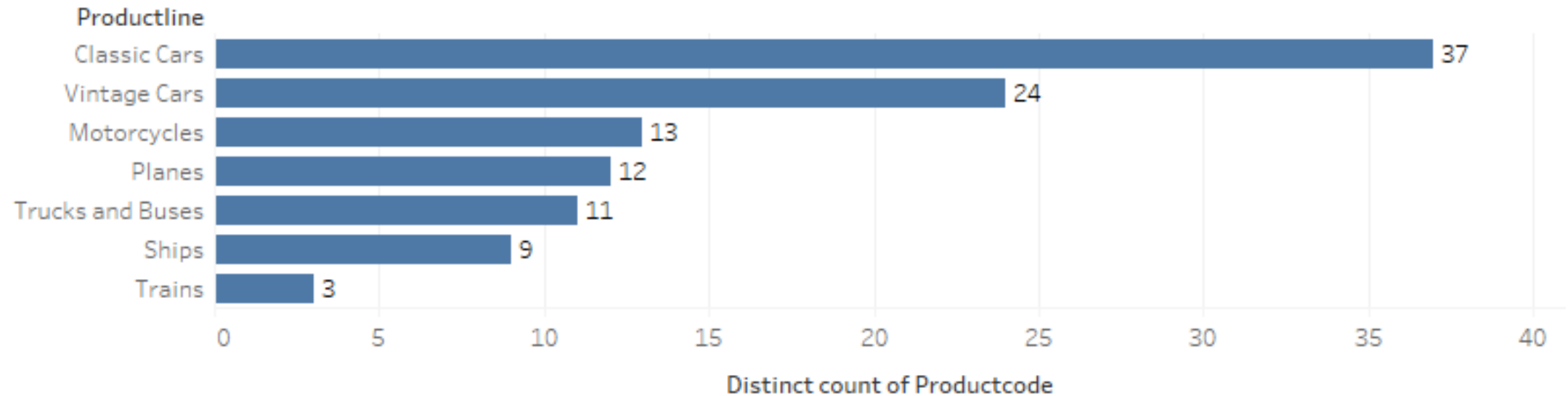


# Y-o-Y Sales in Product lines



The trend of sum of Sales for Orderdate Month broken down by Orderdate Year. Color shows details about Productline.

# Number of Products in each product line



Distinct count of Productcode for each Productline.

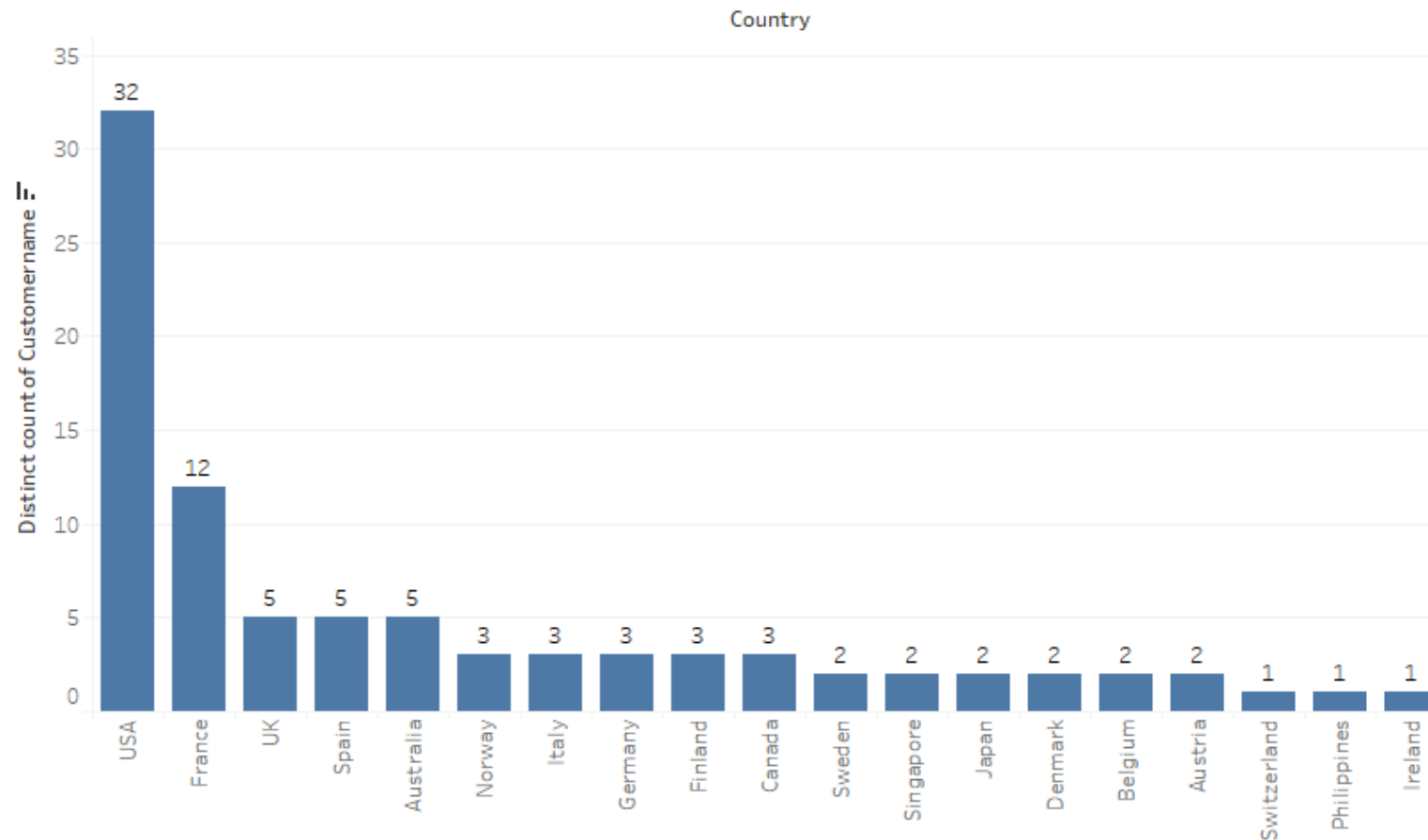
# Table of Quantity sold

Country	Productline / Orderdate																		Year of Orderdate		
	Classic Cars			Motorcycles			Planes			Ships			Trains			Trucks and Buses			Vintage Cars		
	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020
Australia	763	644	411	431	324	121	235	427	151		24	32	33			91	328	286	961	485	499
Austria	248	180	509		197		166	34		113						203			142	80	102
Belgium		30	117					41			343		27	70					20	315	111
Canada	225	147	84		41			317			486					98	253	166	177	264	35
Denmark	478	592	174					70		230	206		38	96			73		157	83	
Finland	280	388	616		301	146	246		175	76		239		89		384			100		152
France	1,210	1,704	626	864	1,093	447	218	321	597	75	455	236		177	45	467	307	293	137	1,575	243
Germany	525	756		43	78			245			55			89			81		142	134	
Ireland		202			58			115						50			37			28	
Italy	304	644				77	460	342	320	71	123		60	22		23	24		562	678	63
Japan		192	122		309			547			127	81			49		102			229	84
Norway	1,001	157		294	190		29	296						72		232	76		199	296	
Philippines	391	87		170	71		215												27		
Singapore	507	536				44				174			174			415	473		182	188	67
Spain	1,072	2,070	1,238	193	260	327	264	465	372	835	332	221	259	112	138	661	505	543	921	1,218	423
Sweden	73	296	183			133	104			287	80			32			433		138	247	
Switzerland		1,078																			
UK	662	845				371	82	397		281	550			168		291			548	818	
USA	4,287	4,805	1,914	1,161	2,768	568	977	1,816	592	702	1,095	460	409	326	177	1,191	2,215	328	2,483	3,094	1,652

Sum of Quantityordered broken down by Productline and Orderdate Year vs. Country. Color shows details about Orderdate Year.



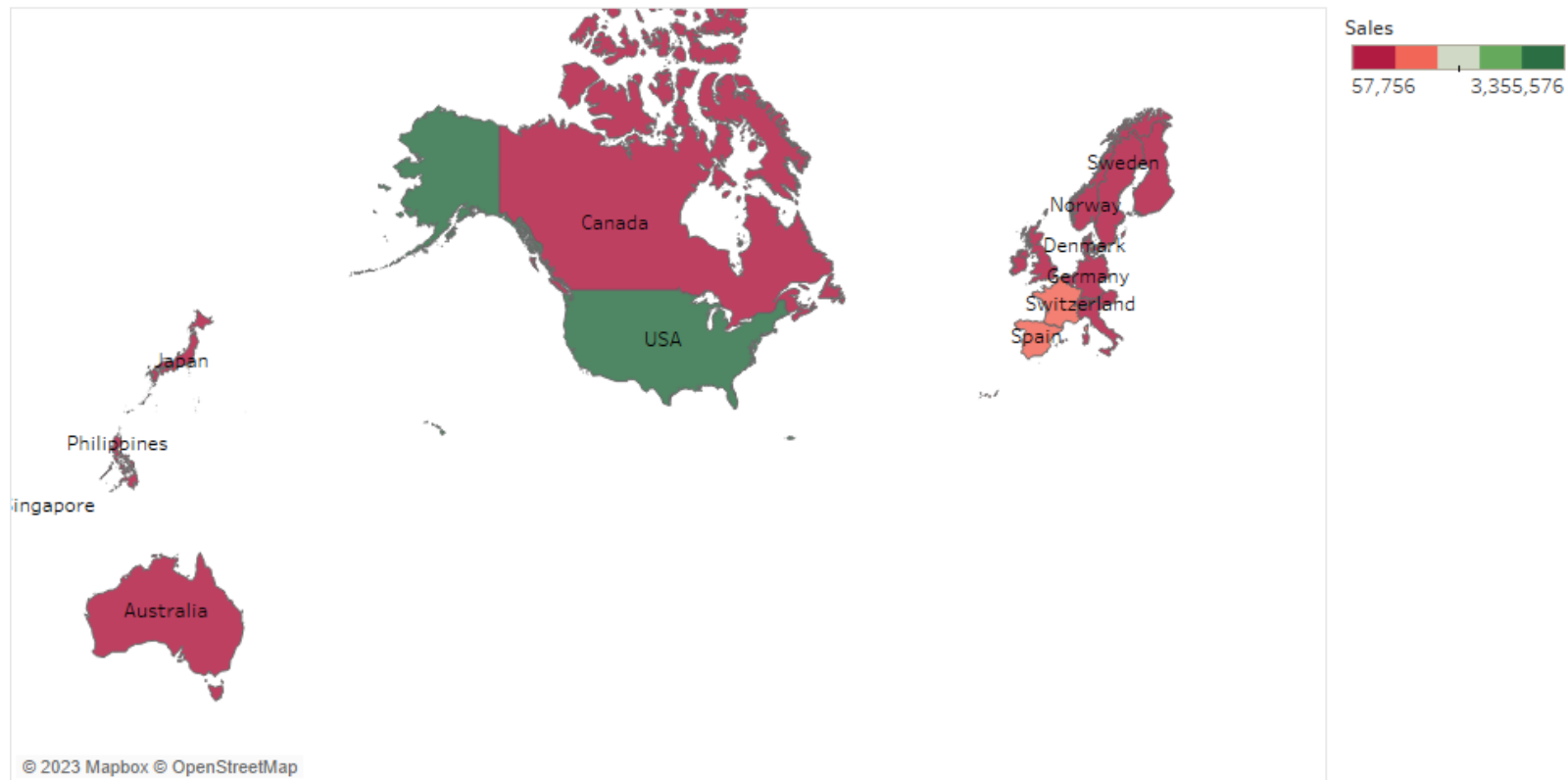
# Customer Base



Distinct count of Customername for each Country.

# Sales Across Geographies

Country Vs Sales



Map based on Longitude (generated) and Latitude (generated). Color shows sum of Sales. The marks are labeled by Country.



# Pivot table for RFM analysis – customer count

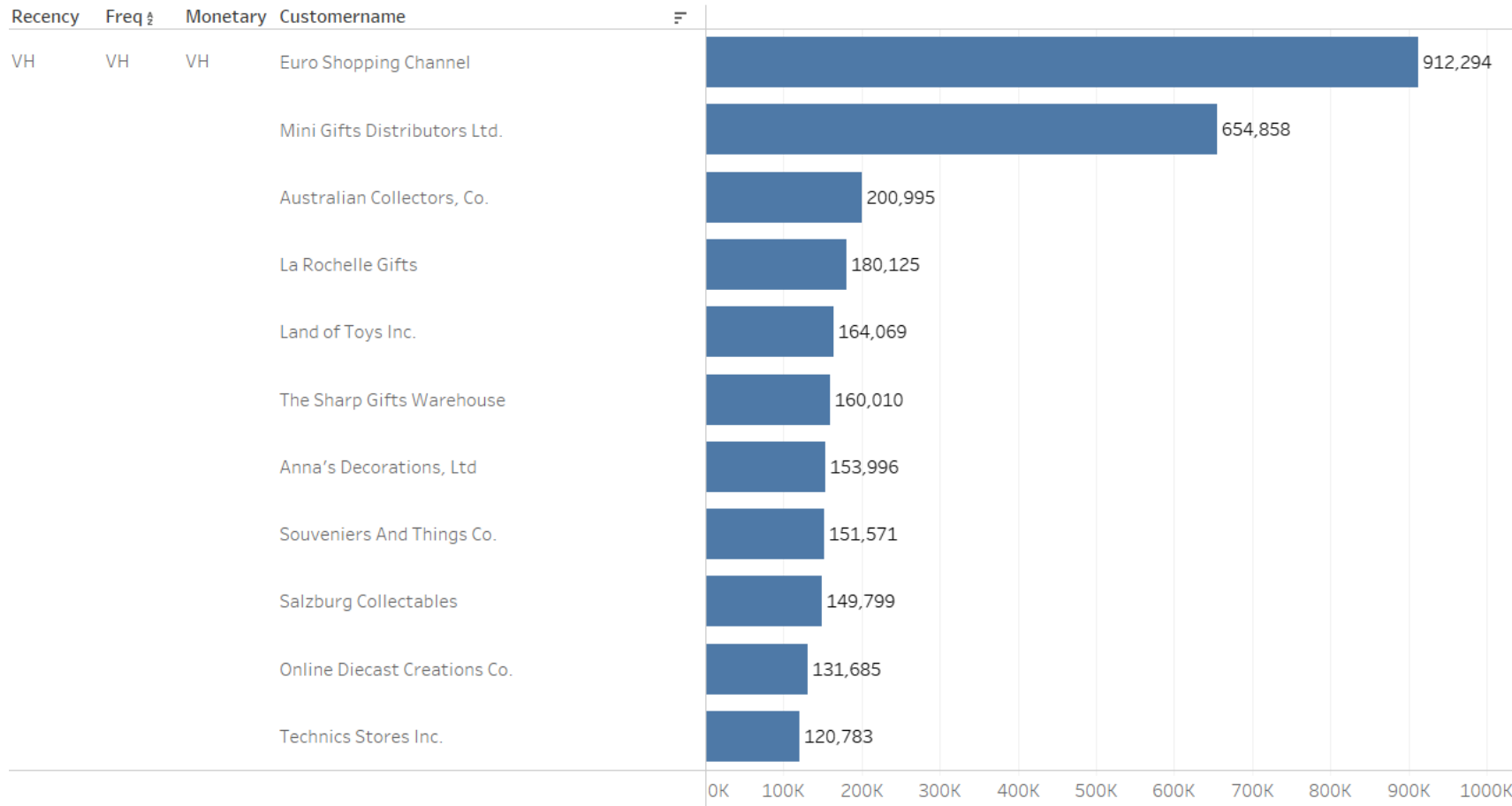
	Recency	Freq	Monetary					
			VH	H	M	L	Grand Total	
Very Active	VH	VH	11				11	Active and loyal
		H	1	4			5	Alive but not too active but good value
		M		2	3		5	Alive but less value
		L			1	1	2	
Active	H	VH	2				2	Active and loyal
		H	1	5			6	Alive but not too active but good value
		M		3	5	1	9	Alive but less value
		L			2	3	5	
At risk	M	VH	5	2			7	High value but havent purchased lately
		H		2			2	medium value and less frequent
		M		1	3		4	low value,less frequent and less recent
		L			2	7	9	
Lost	L	VH	2				2	High value but not recent
		H		1	1		2	medium value, not frequent and not recent
		M		2	4	1	7	low value,less frequent and less recent
		L			1	10	11	
	Grand Total		22	22	22	23	89	

# Pivot table for RFM analysis – Sales percentage

	Row Labels		VH	H	M	L	Grand Total					
Very Active	VH	VH	31%	0%	0%	0%	31%		Active and loyal			
		H	1%	4%	0%	0%	6%		Alive but not too active but good value			
		M	0%	2%	2%	0%	5%		Alive but less value			
		L	0%	0%	1%	1%	1%					
Active	H	VH	3%	0%	0%	0%	3%		Active and loyal			
		H	1%	6%	0%	0%	7%		Alive but not too active but good value			
		M	0%	3%	4%	1%	8%		Alive but less value			
		L	0%	0%	2%	1%	3%					
At risk	M	VH	8%	2%	0%	0%	10%		High value but havent purchased lately			
		H	0%	2%	0%	0%	2%		medium value and less frequent			
		M	0%	1%	3%	0%	3%		low value,less frequent and less recent			
		L	0%	0%	2%	4%	5%					
Lost	L	VH	3%	0%	0%	0%	3%		High value but not recent			
		H	0%	1%	1%	0%	2%		medium value, not frequent and not recent			
		M	0%	2%	3%	1%	6%		low value,less frequent and less recent			
		L	0%	0%	1%	5%	6%					
	Grand Total			47%	24%	18%	11%	100%				

# Best Customers

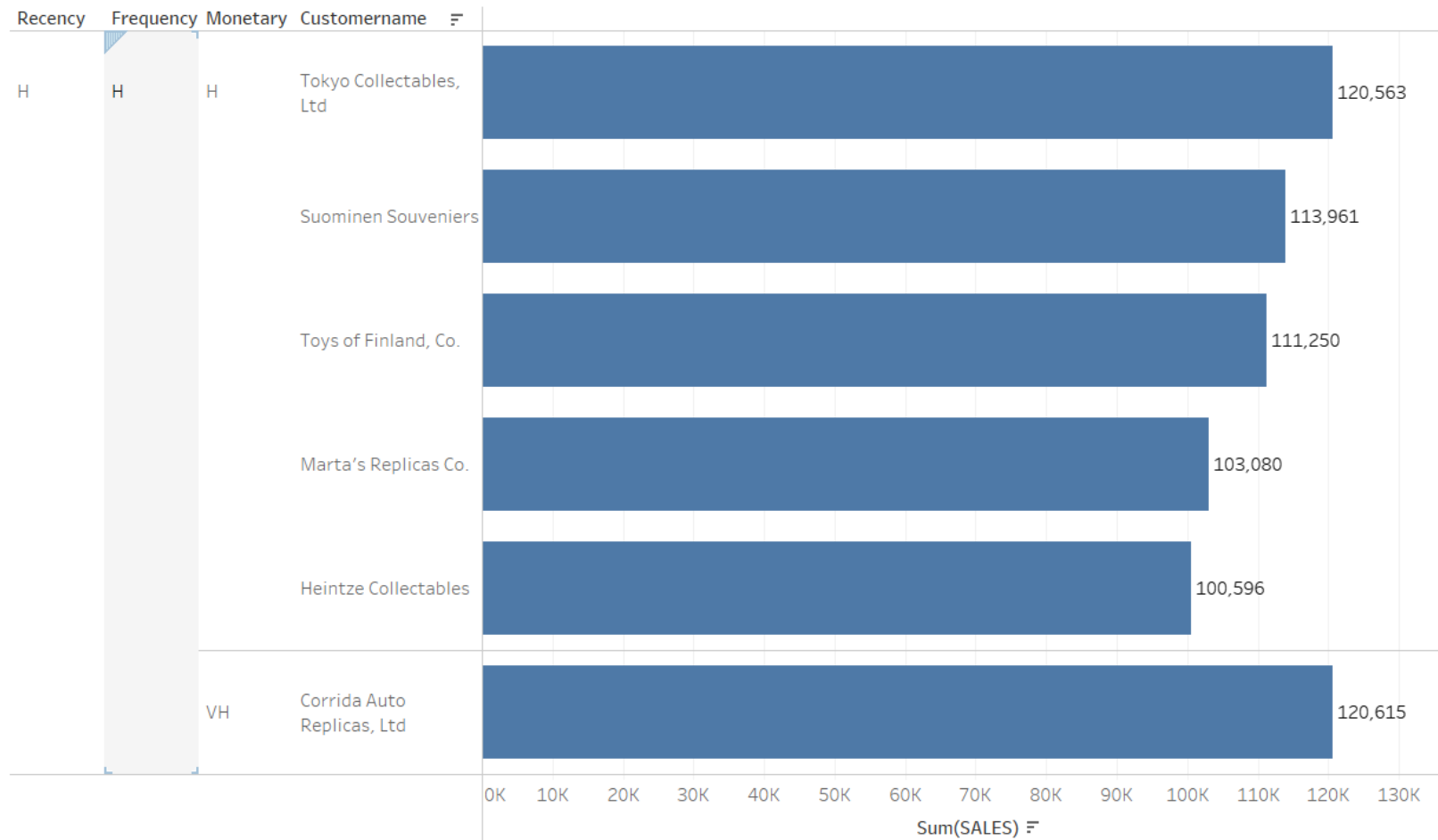
High valuable customers with very high recency and frequency



- Best customers with
- Very High recency
  - Very High frequency
  - and
  - Very high monetary

# Loyal Customers

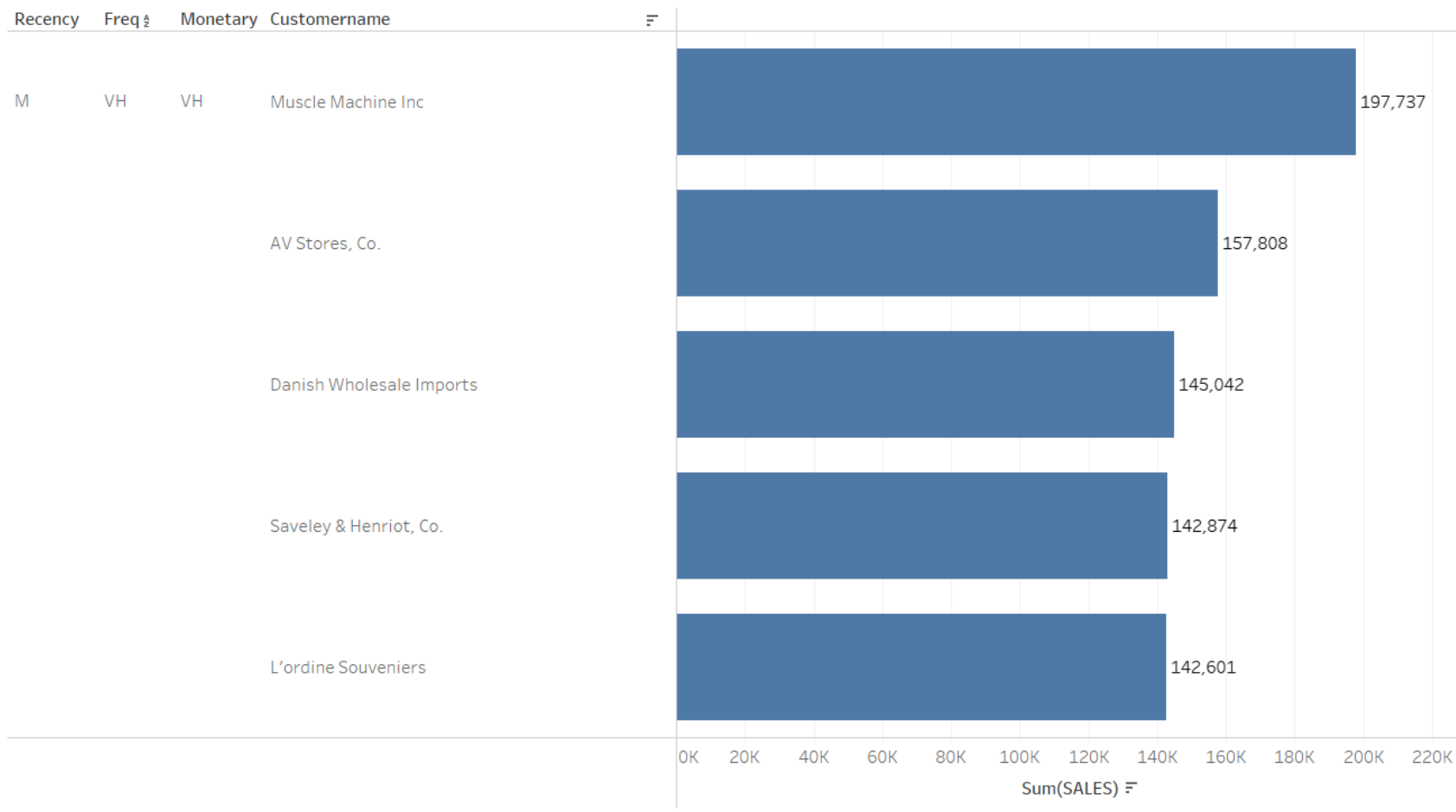
High value Customers who are more frequent and recent apart from best customers



- Loyal customers with
- High Recency
  - High frequency and
  - Very high or High monetary

# Customers in the verge of churning

Medium recent customers with very high frequency

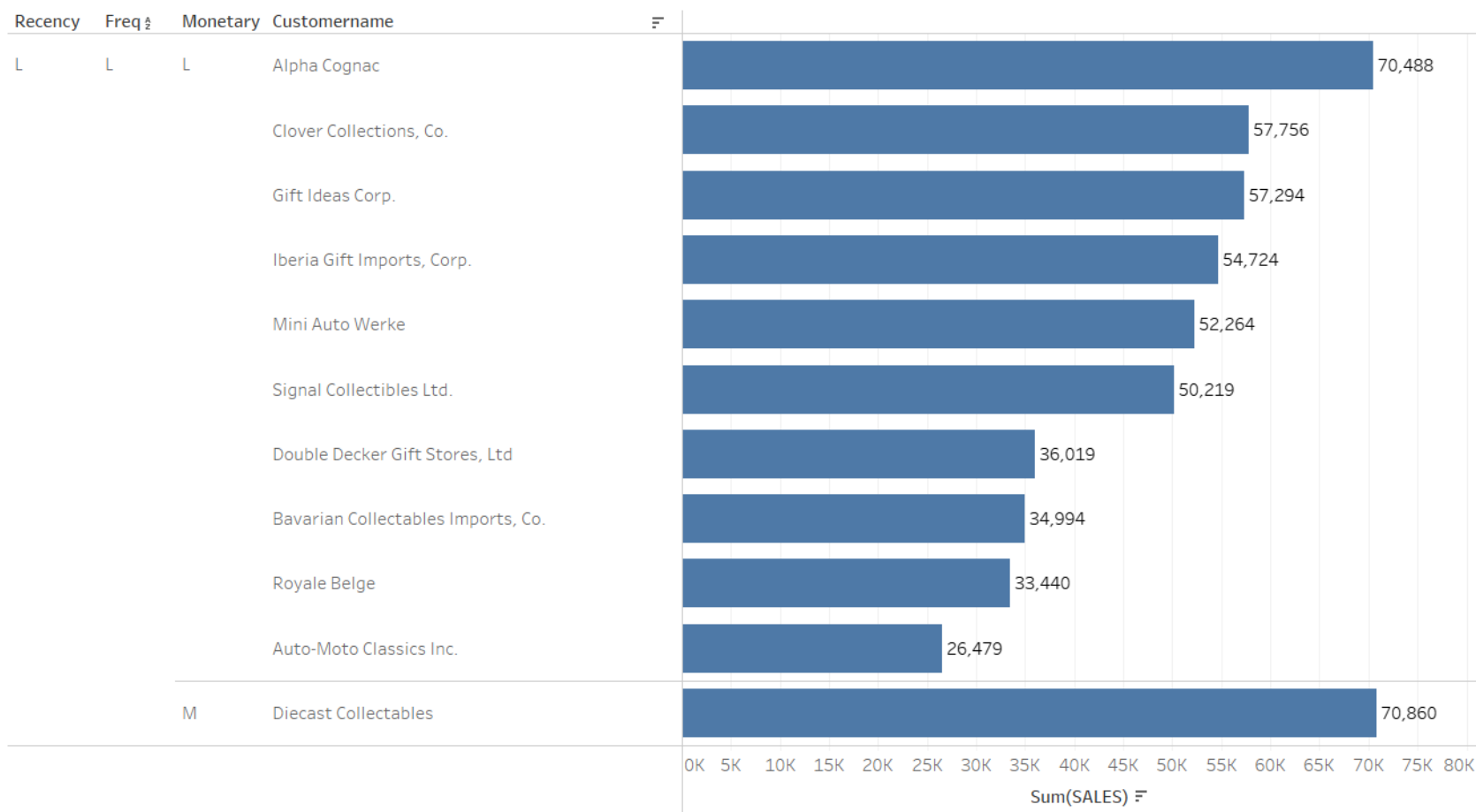


- At Risk customers with
- Medium Recency
  - Very High frequency
  - and
  - Very high monetary



# Lost Customers

Low value customers with very less frequency and recency



Lost customers with

- Low Recency
- Low frequency and
- Low monetary