



NYC DATA SCIENCE
ACADEMY

Choosing an effective visualization

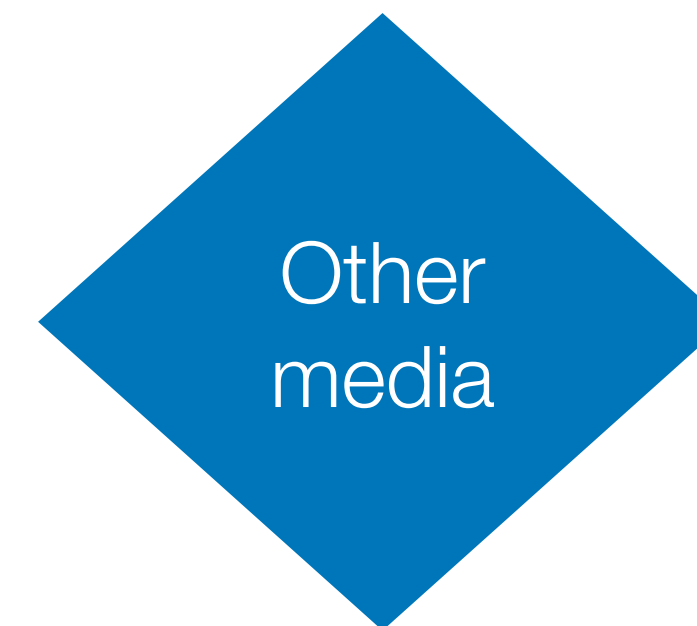
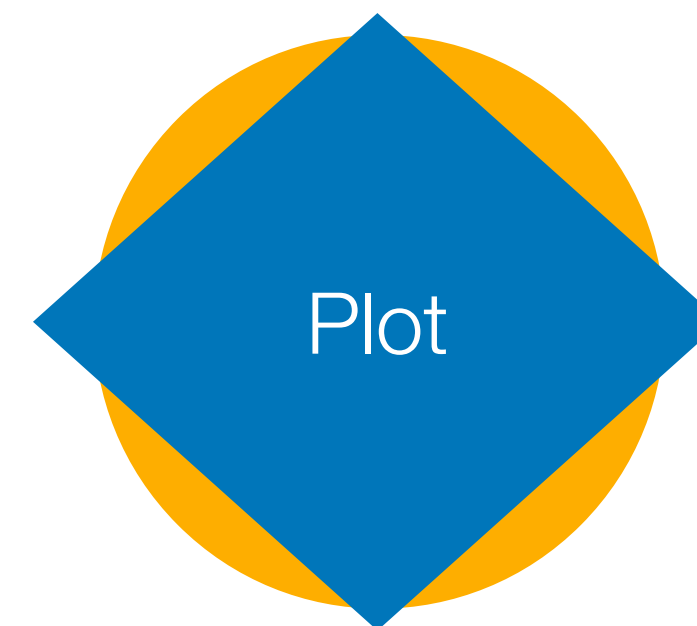
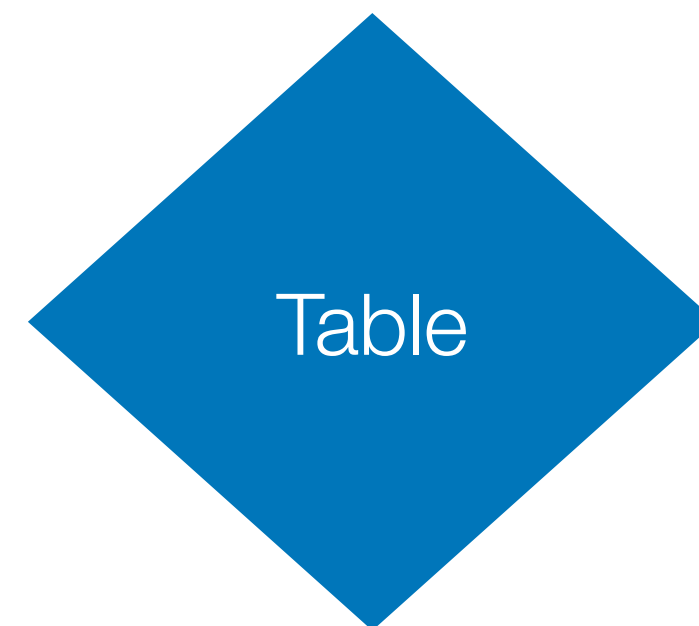
Choosing an effective visualization

Presenting your results

1) Question

2) Relevant information

3) Target audience



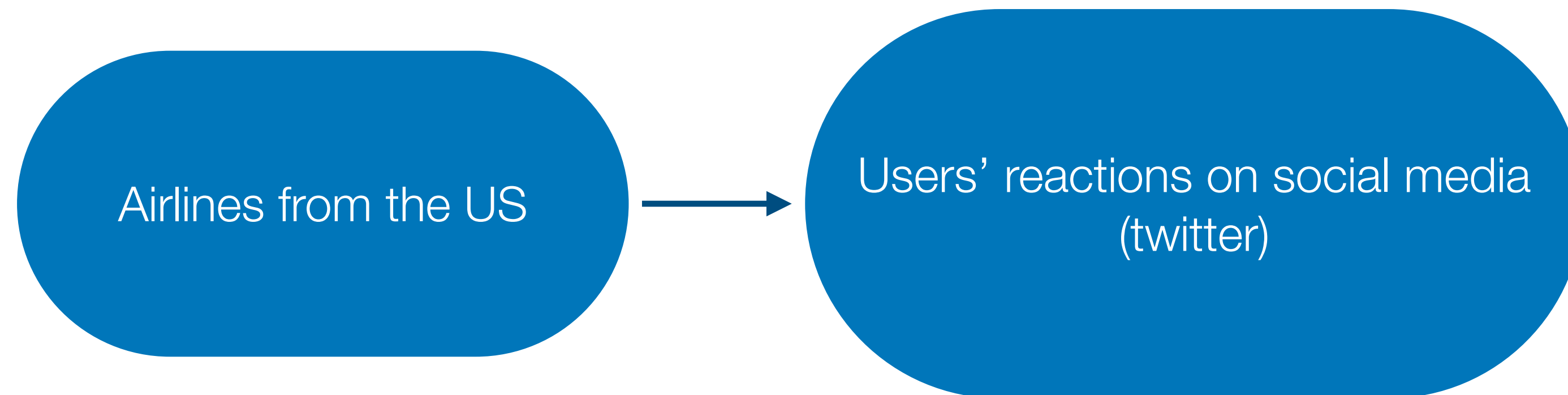
Negative sentiment on Airlines

1) We are a consulting company and want to sell a program to improve user experience to the airlines.

What is the users satisfaction with the airlines in the US?

2) NLP of users' reactions on social media

3) Business team from American Airlines



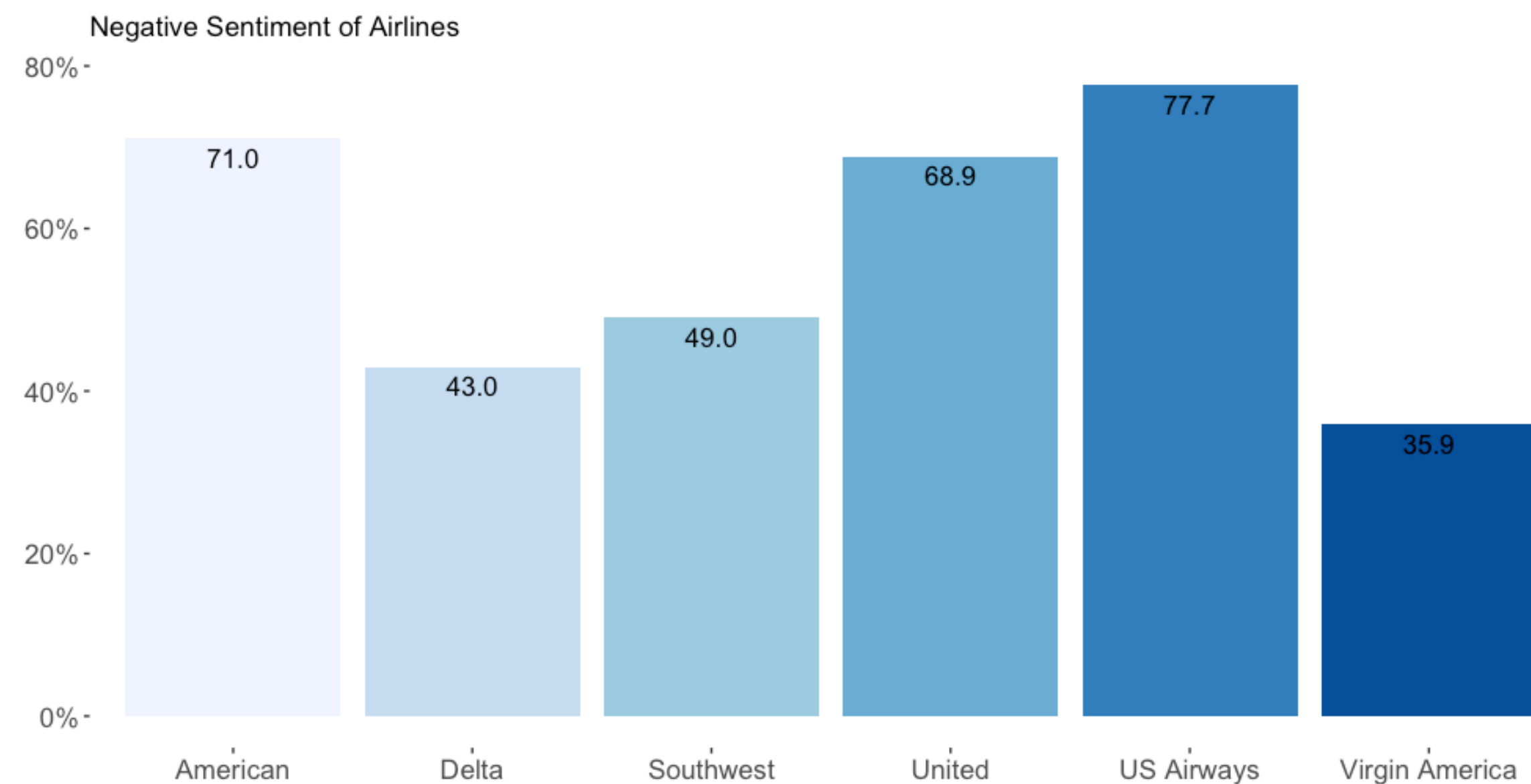
Negative sentiment on Airlines

	tweet_id	airline_sentiment	airline_sentiment_confidence	negativereason	negativereason_confidence	airline	airline_sentiment_gold	name	negativereason_gold	retweet_count	text
8966	5.675883e+17	neutral	1.0000		NA	Delta		JetBlueNews		0	@JetBlue's new CEO seeks the righ
8965	5.675900e+17	negative	1.0000	Can't Tell	0.6503	Delta		nesi_1992		0	@JetBlue is REALLY getting on my
4326	5.675915e+17	negative	1.0000	Late Flight	0.3460	United		CPoutloud		0	@united yes. We waited in line for
4325	5.675924e+17	negative	1.0000	Late Flight	1.0000	United		brenduch		0	@united the we got into the gate a
6746	5.675944e+17	negative	1.0000	Customer Service Issue	0.3451	Southwest		VahidESQ		0	@SouthwestAir its cool that my ba
4324	5.675946e+17	negative	1.0000	Bad Flight	0.6707	United		brenduch		0	@united and don't hope for me ha
4323	5.675957e+17	negative	1.0000	Late Flight	1.0000	United		CRomerDome		0	@united I like delays less than you
4322	5.676140e+17	negative	1.0000	Customer Service Issue	0.3545	United		JustOGG		0	@united, link to current status of
6745	5.676171e+17	negative	1.0000	Customer Service Issue	1.0000	Southwest		mrshossruns		0	@SouthwestAir you guys there? Ar
4321	5.676175e+17	negative	1.0000	Customer Service Issue	0.6797	United		feliciastoler		0	@united I tried 2 DM it would not
4320	5.676232e+17	negative	0.6337	Flight Booking Problems	0.6337	United		markhlyon		0	Wanted to get my bag benefit, but
4319	5.676273e+17	negative	1.0000	Bad Flight	0.3611	United		DBsViewOnThings		0	@united please see a flight attend
4318	5.676303e+17	negative	1.0000	Customer Service Issue	0.6752	United		ljsbrooks		0	@united still waiting for a reply
4317	5.676341e+17	neutral	1.0000		NA	United		gwaki		0	@united even though technically a
11875	5.676433e+17	neutral	1.0000		NA	US Airways		ashenfaced		0	@USAirways how's us 1797 lookin
6744	5.676555e+17	positive	1.0000		NA	Southwest		rjp1208		0	@SouthwestAir nice work on the u

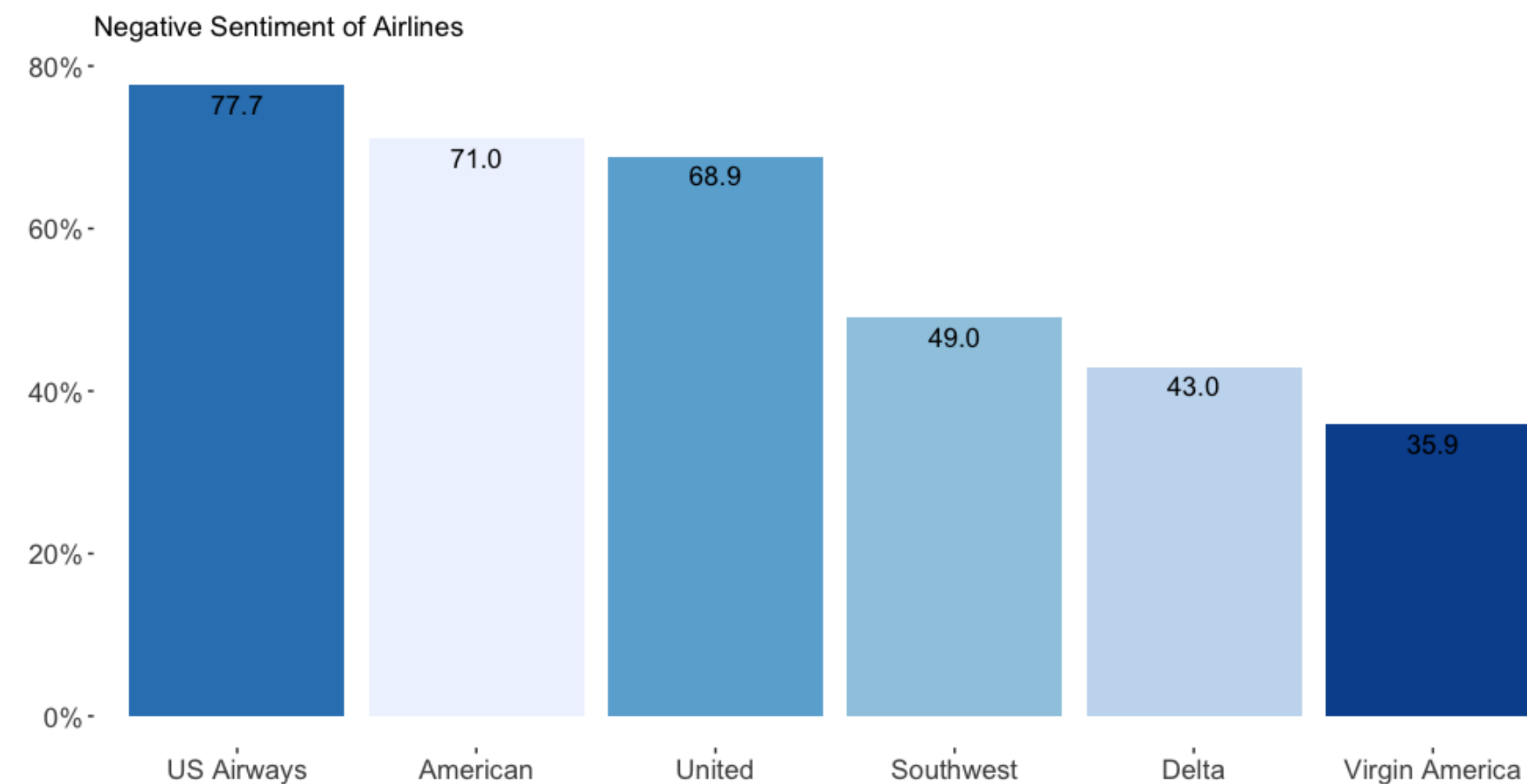
airline	n	percentage of negative tweets
American	2759	71.04
Delta	2222	42.98
Southwest	2420	49.01
United	3822	68.89
US Airways	2913	77.69
Virgin America	504	35.91

Negative sentiment on Airlines

1



2



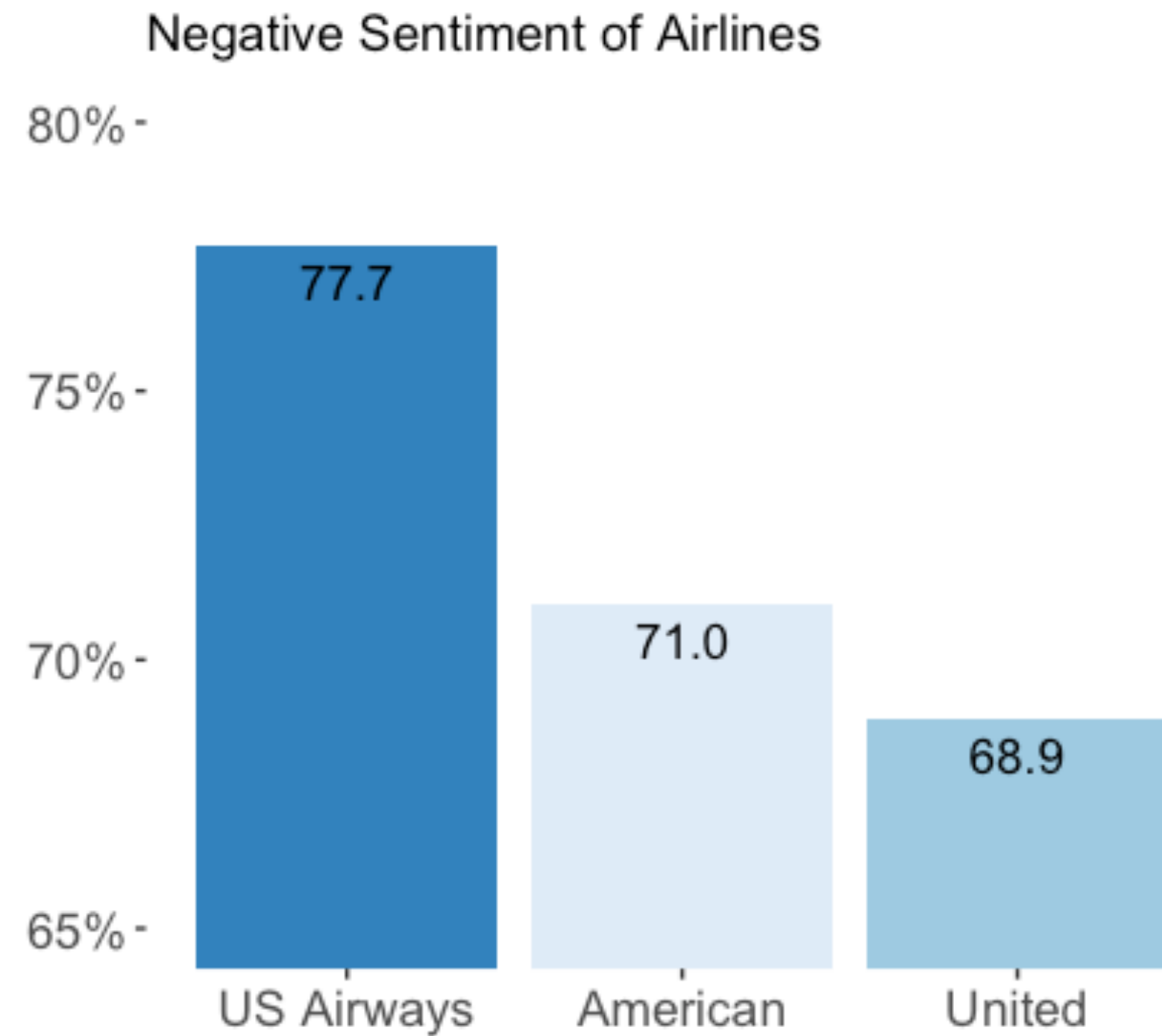
Bar Chart

- Sorting categories gives the viewer a sense of order

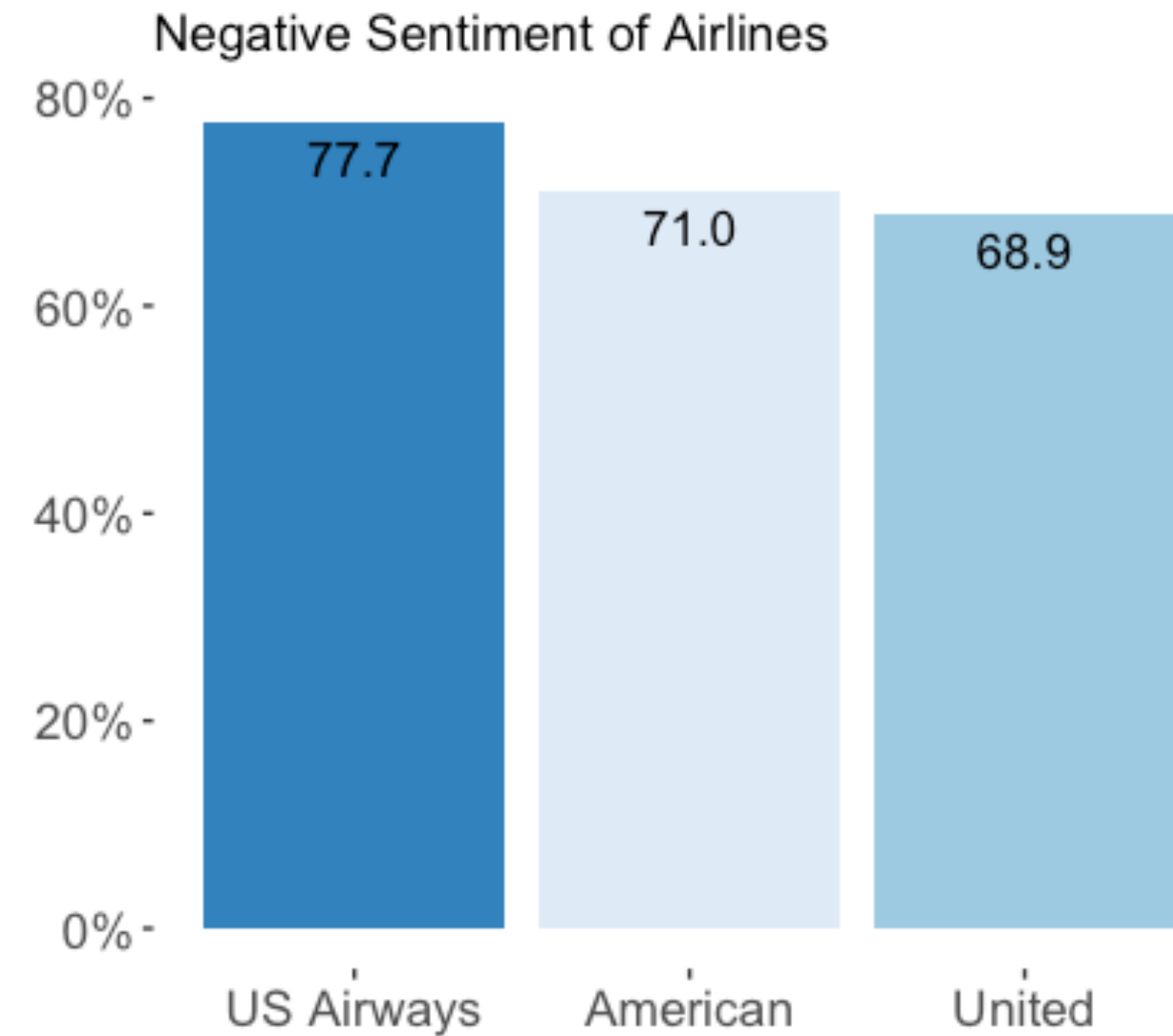
Bar Chart

In detail

- Because of the way our eyes compare the relative end points of the bars, it's important to have the context of the entire bar there in order to make a fair comparison
- If the business problem involves the specific differences, adjusting the axis can aid to make a point



1



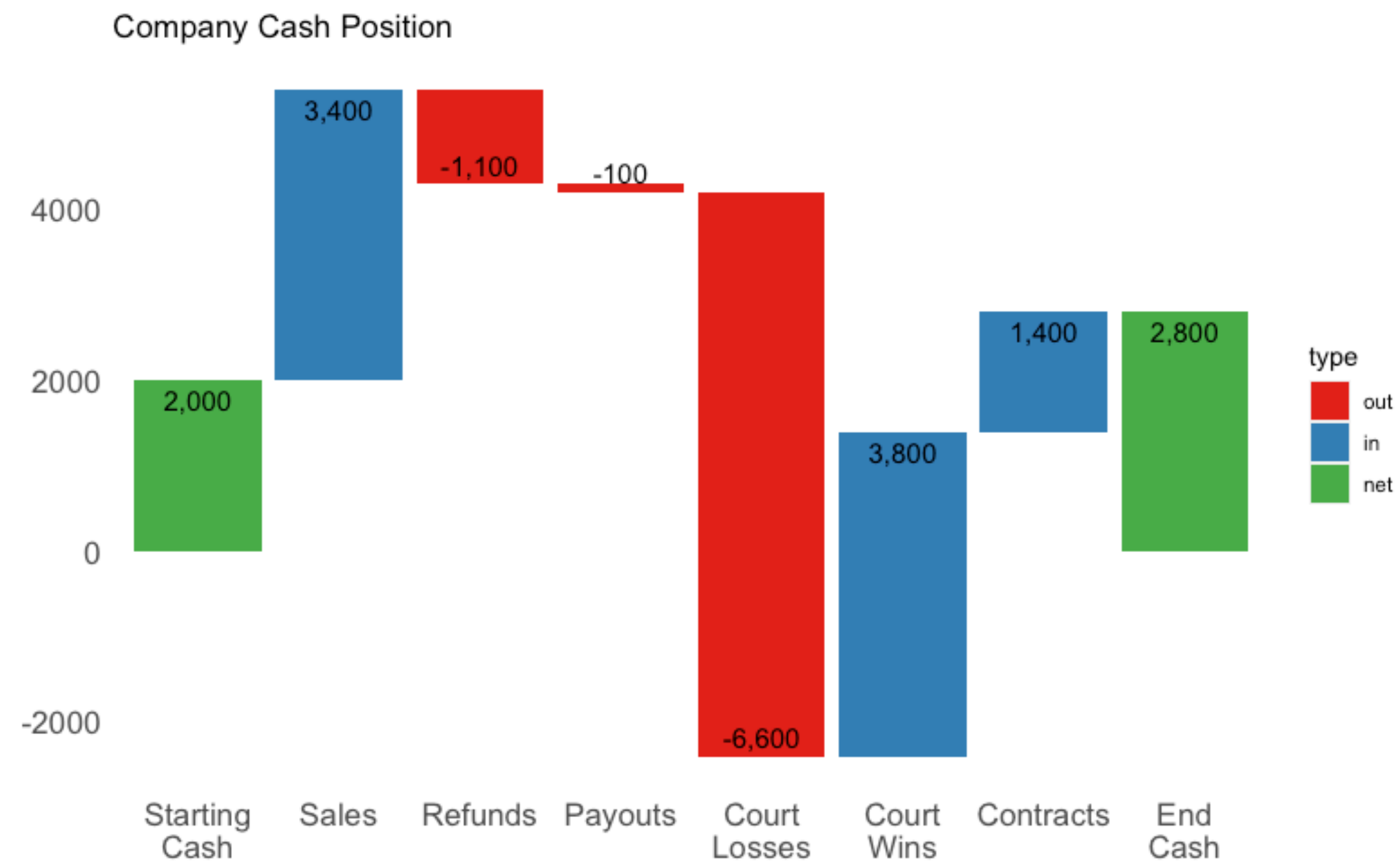
2

Law firm finances

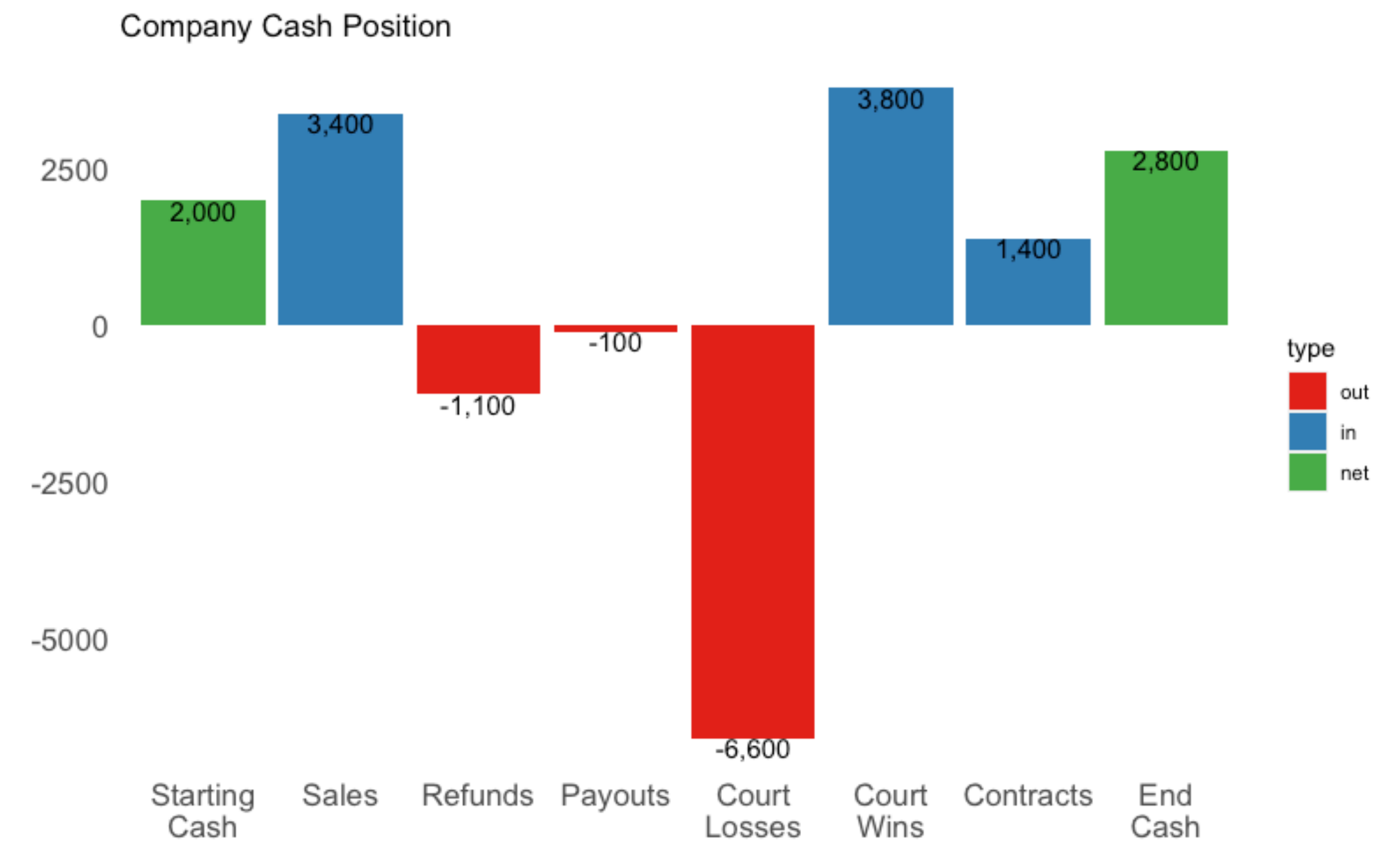
- 1) Impact of each transaction category on the cash position of the Company, to identify areas of improvement
- 2) Summary of the company’s transactions
- 3) Multidisciplinary team at the company

	description	type	\$ amount
1	Starting Cash	net	2000
2	Sales	in	3400
3	Refunds	out	−1100
4	Payouts	out	−100
5	Court Losses	out	−6600
6	Court Wins	in	3800
7	Contracts	in	1400
8	End Cash	net	2800

1



2



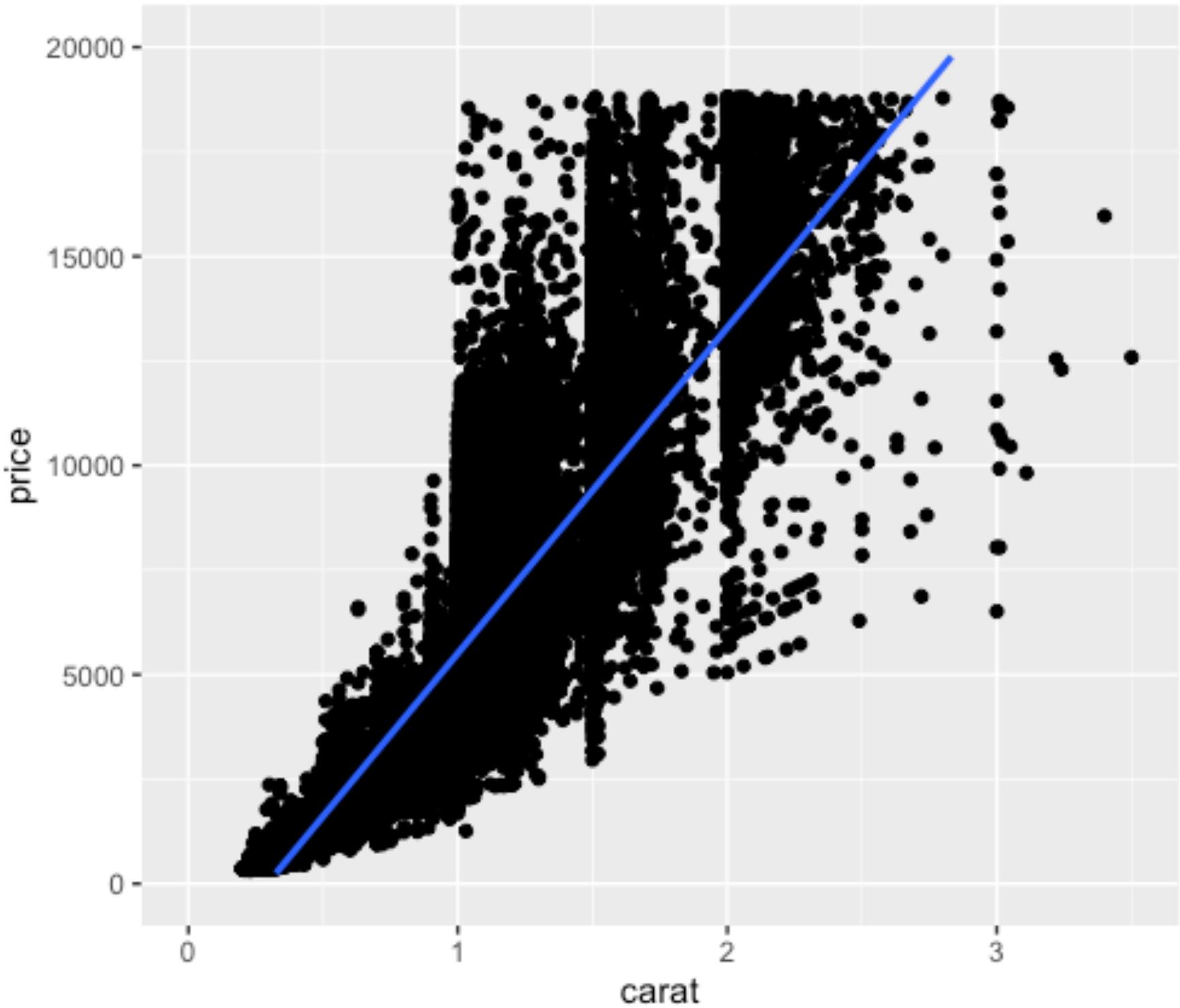
Waterfall chart

- A waterfall chart shows the cash flow
- The final column represents the end cash after all the additions and deductions have been applied to the starting cash

Diamond market

- 1) We want to add diamonds too our inventory (small investment and big profit)
- 2) Database of diamond features and prices
- 3) Company’s CEO

Variable	Description	Values
price	price in US dollars	\$326-\$18,823
carat	weight of the diamond	0.2-5.01
cut	quality of the cut	Fair, Good, Very Good, Premium, Ideal

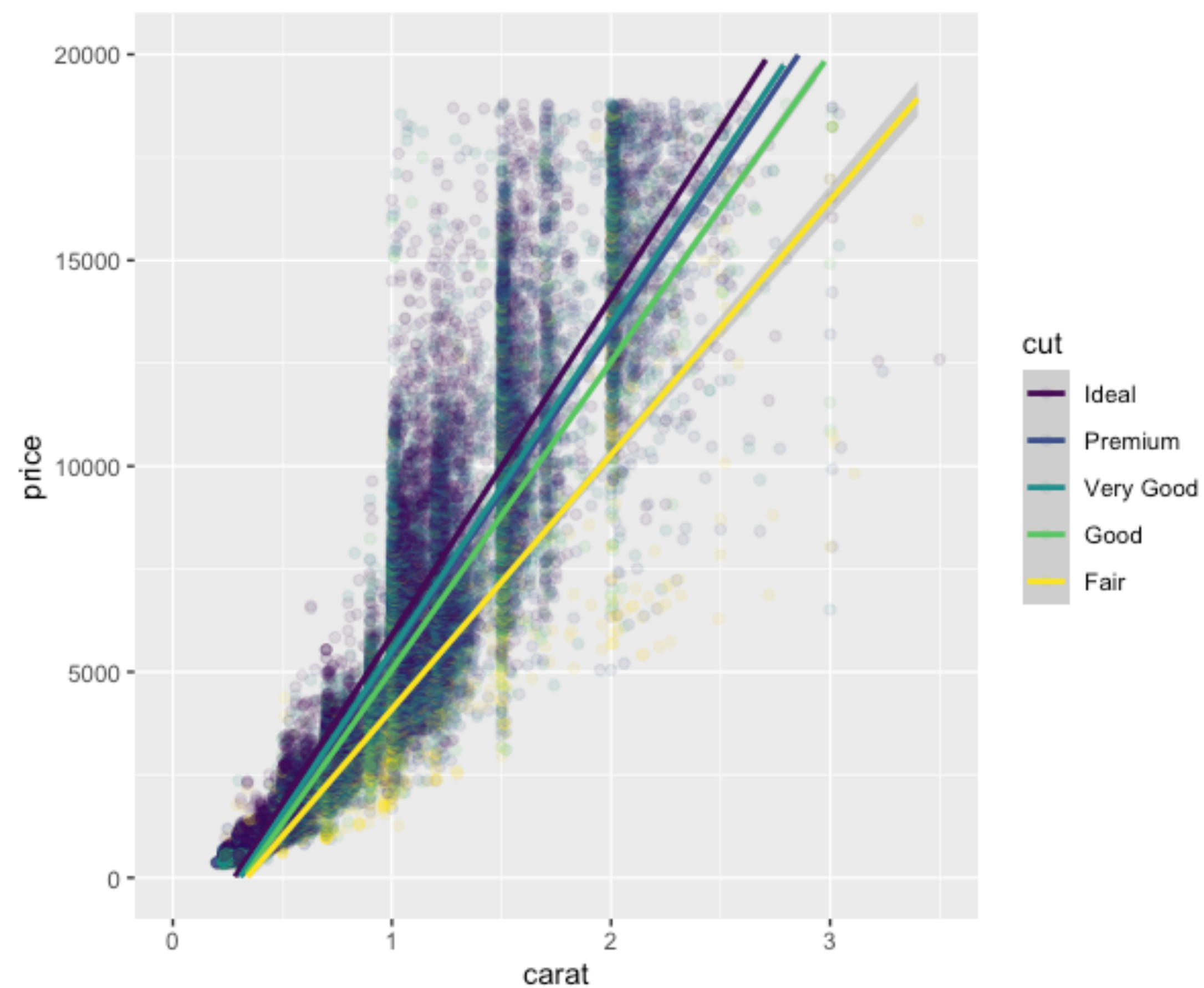
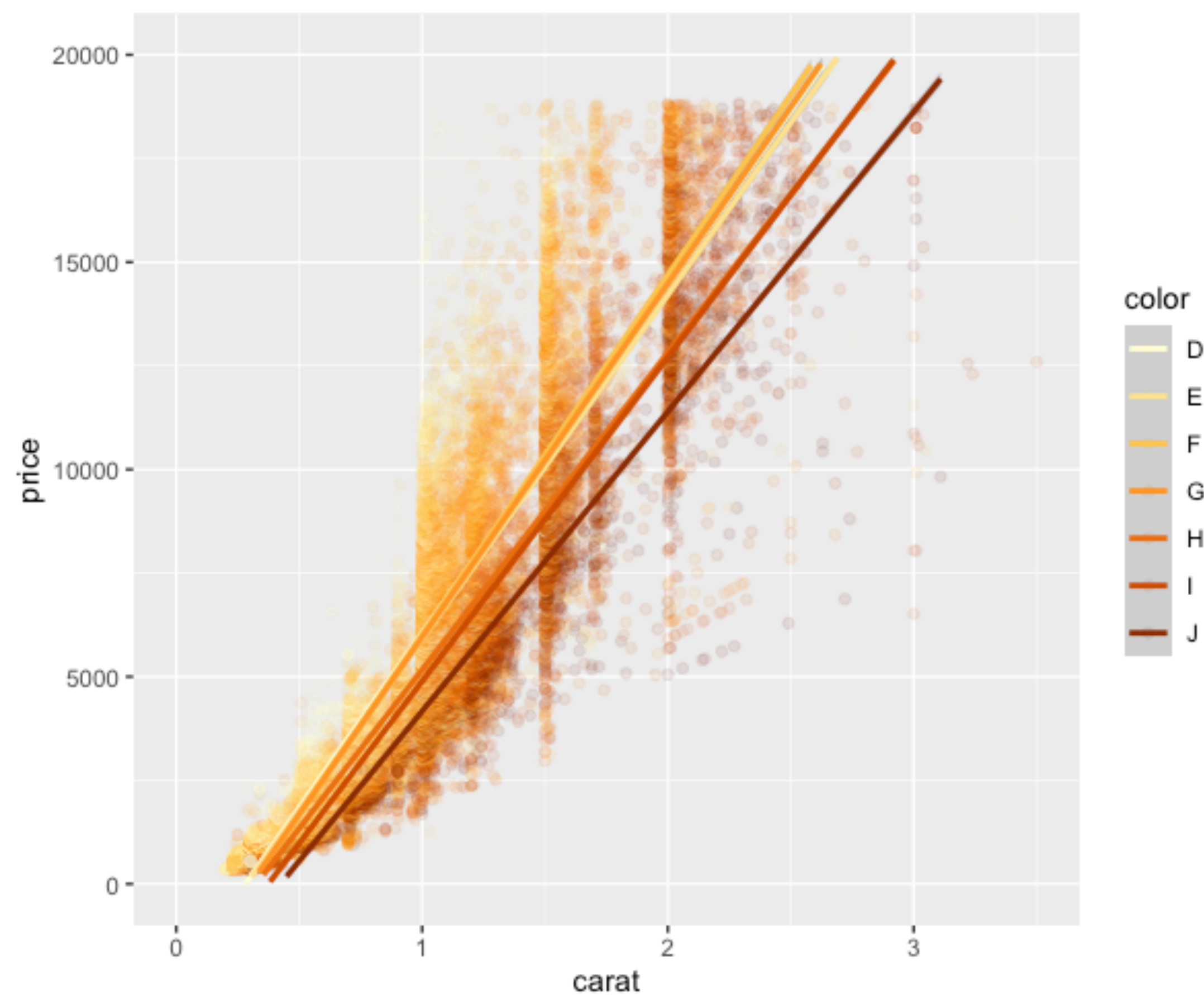


- Modify the ranges of x and y shown in the plot to “zoom-in”

Scatter plot

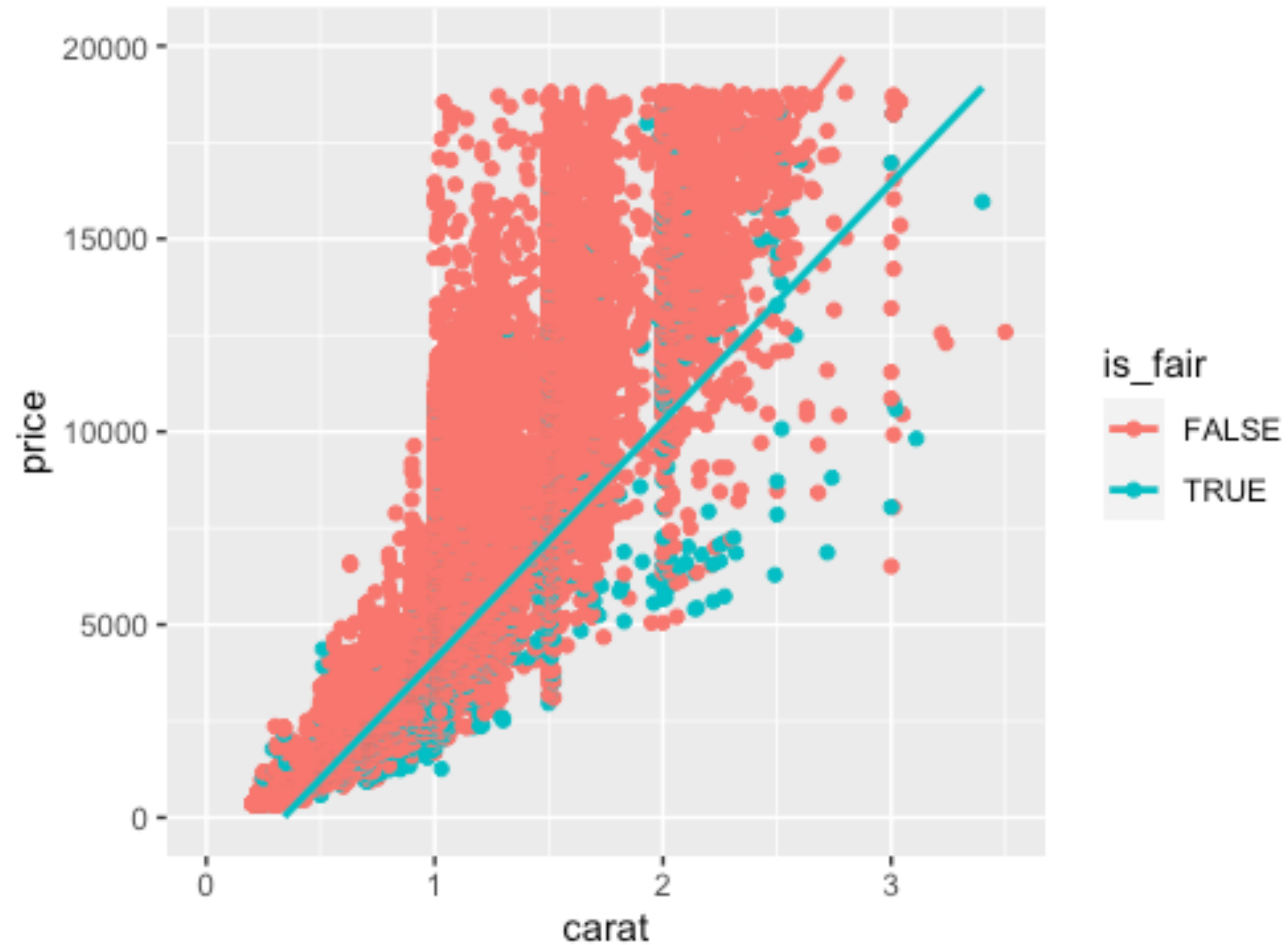
Diamond market

Adding a categorical feature to get insights

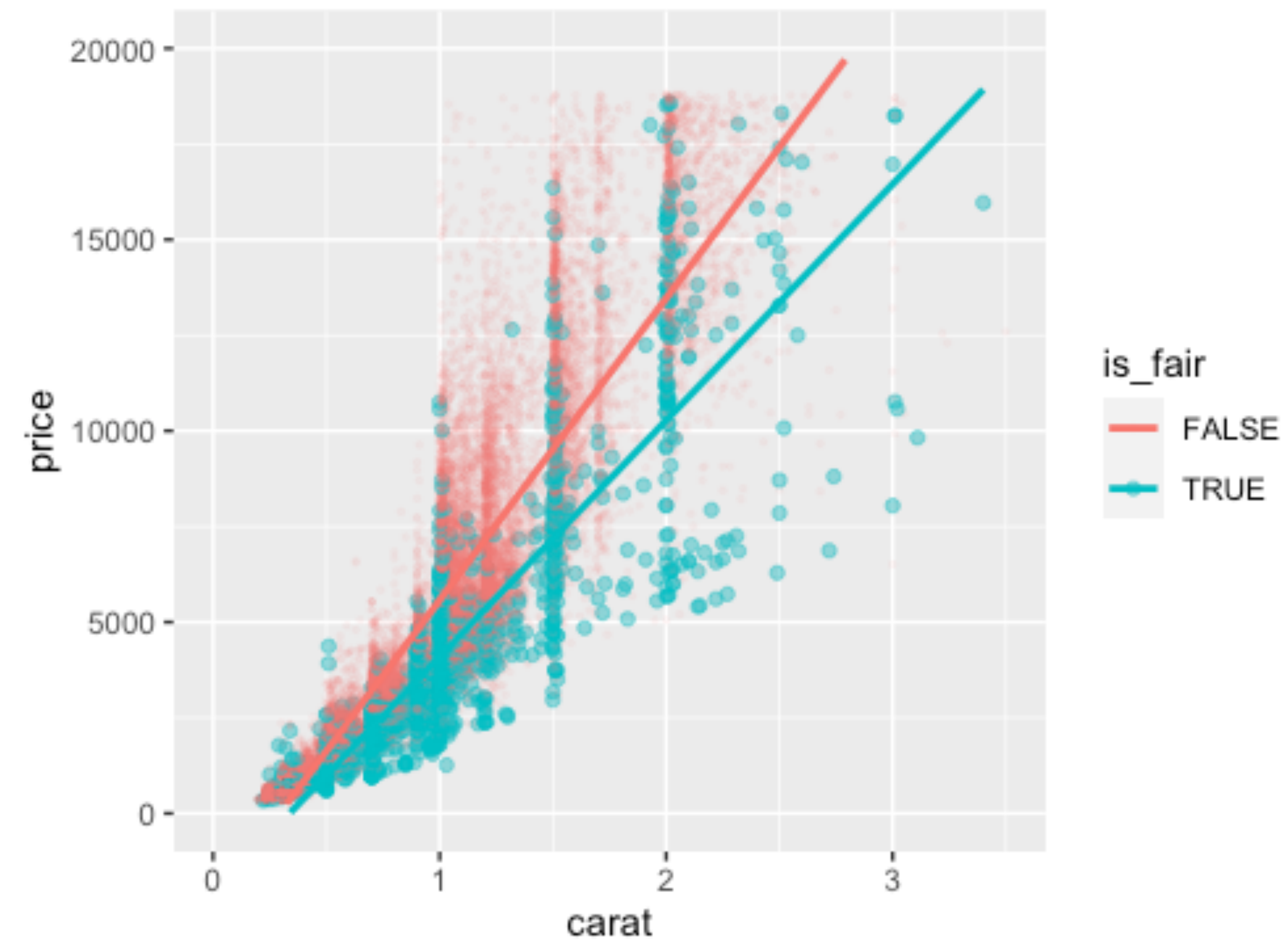


Diamond market

Combining the categories



Adjusting transparency and size



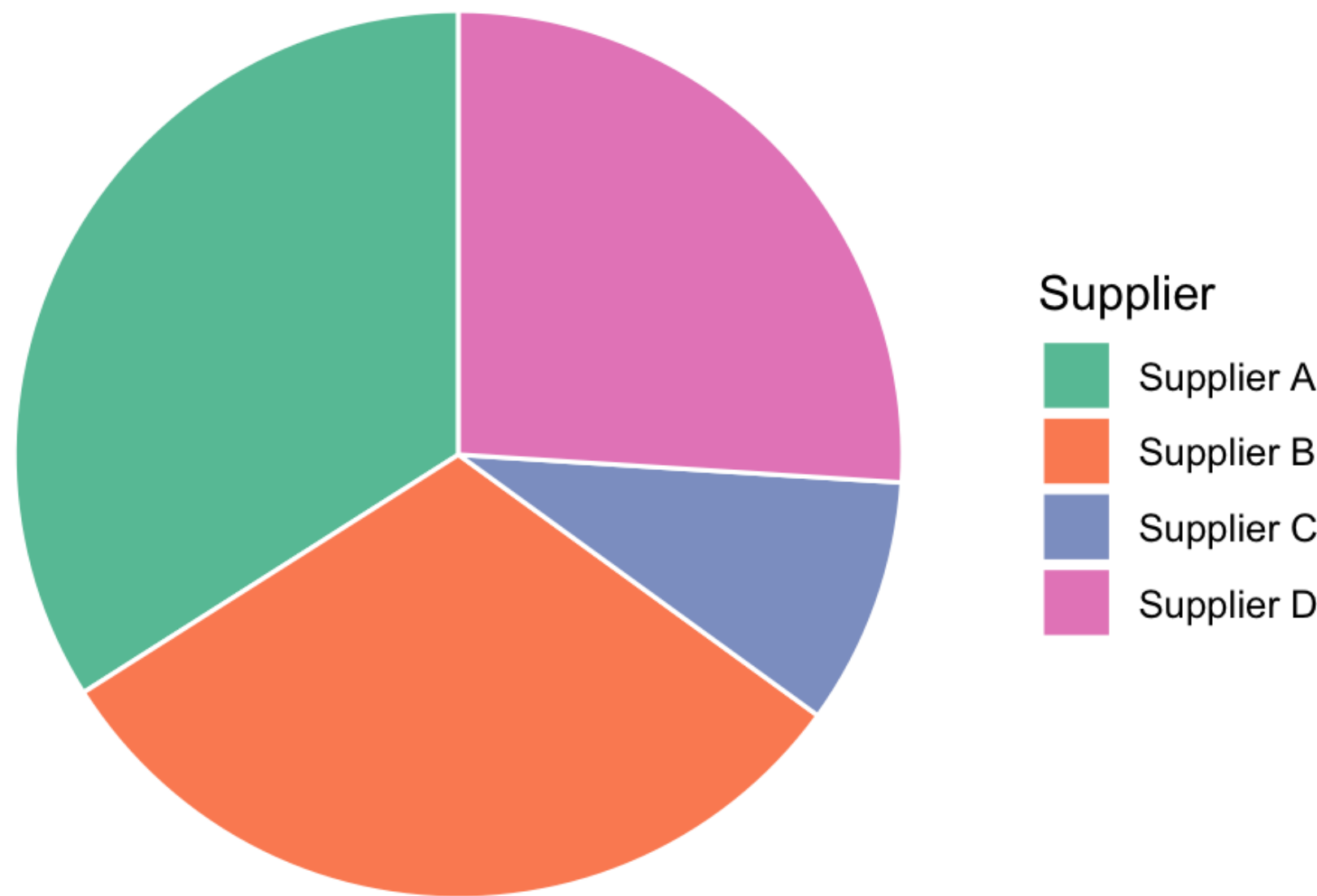
- Avoids overlapping, highlights the 'fair' category

Supplier market share

- 1) We want to identify the largest supplier of our source materials
- 2) Summary of the supplier market share
- 3) Multidisciplinary team

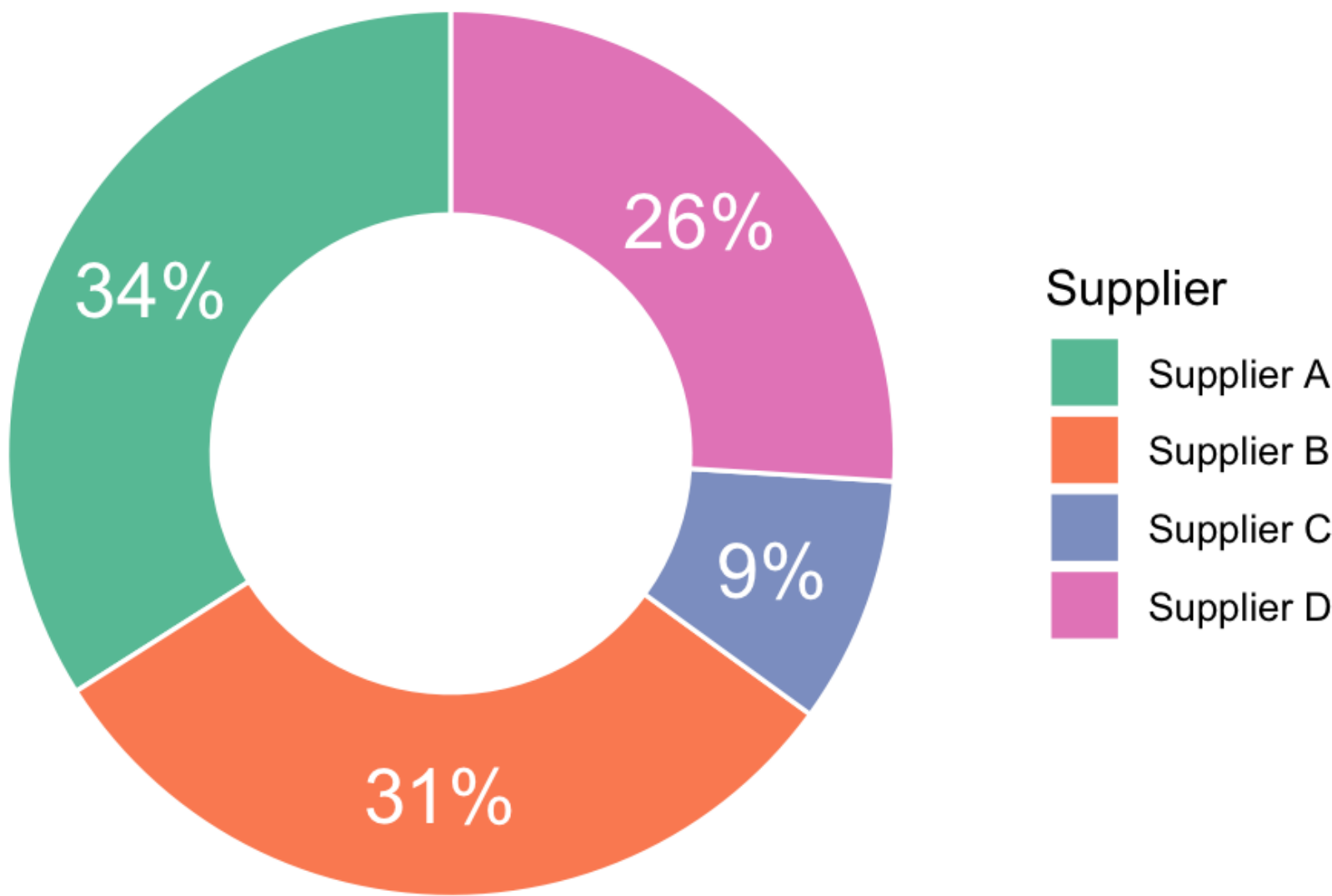
	Supplier	prop
1	Supplier A	34
2	Supplier B	31
3	Supplier C	9
4	Supplier D	26

Supplier Market Share



1

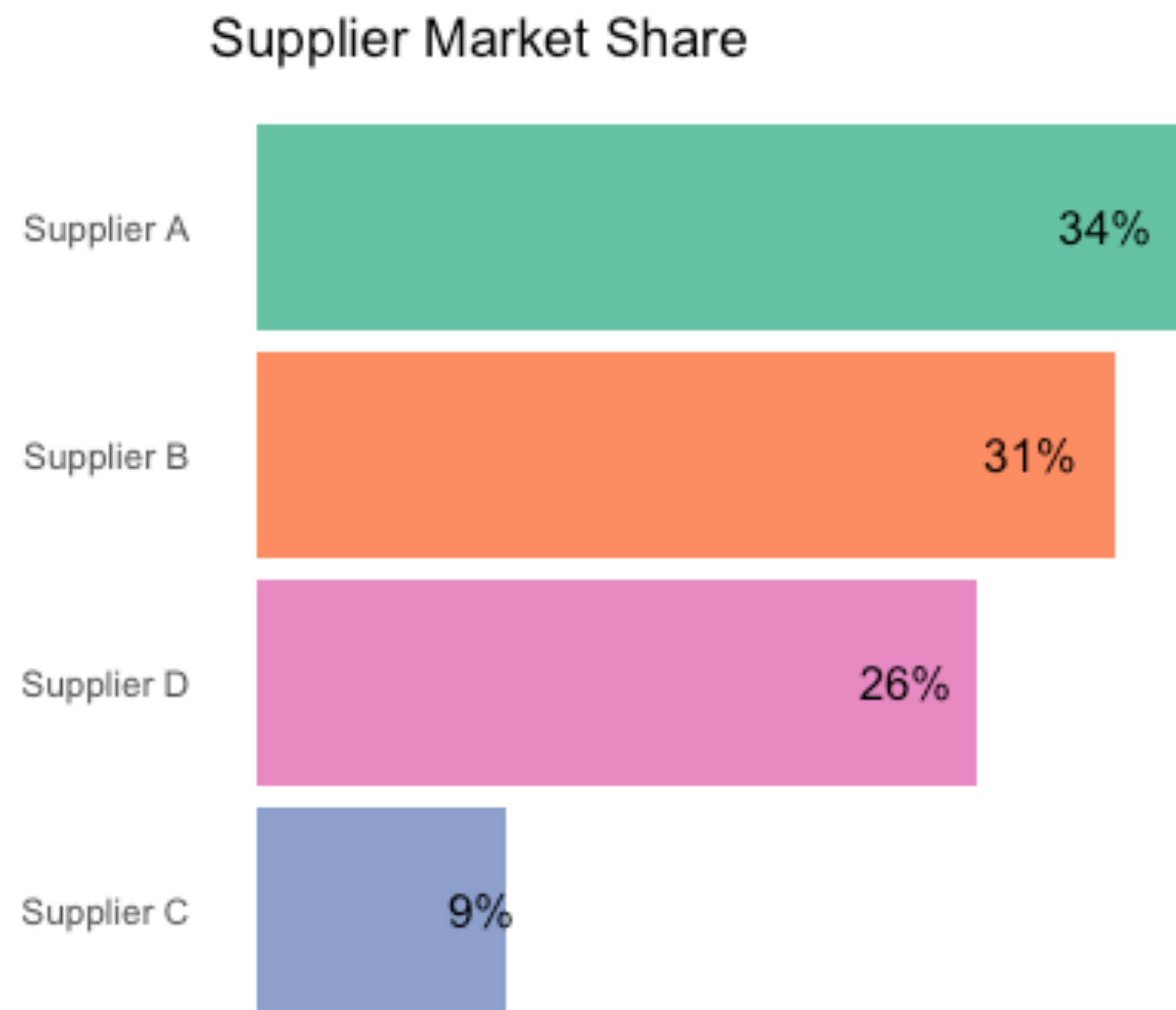
Supplier Market Share



2

Supplier market share

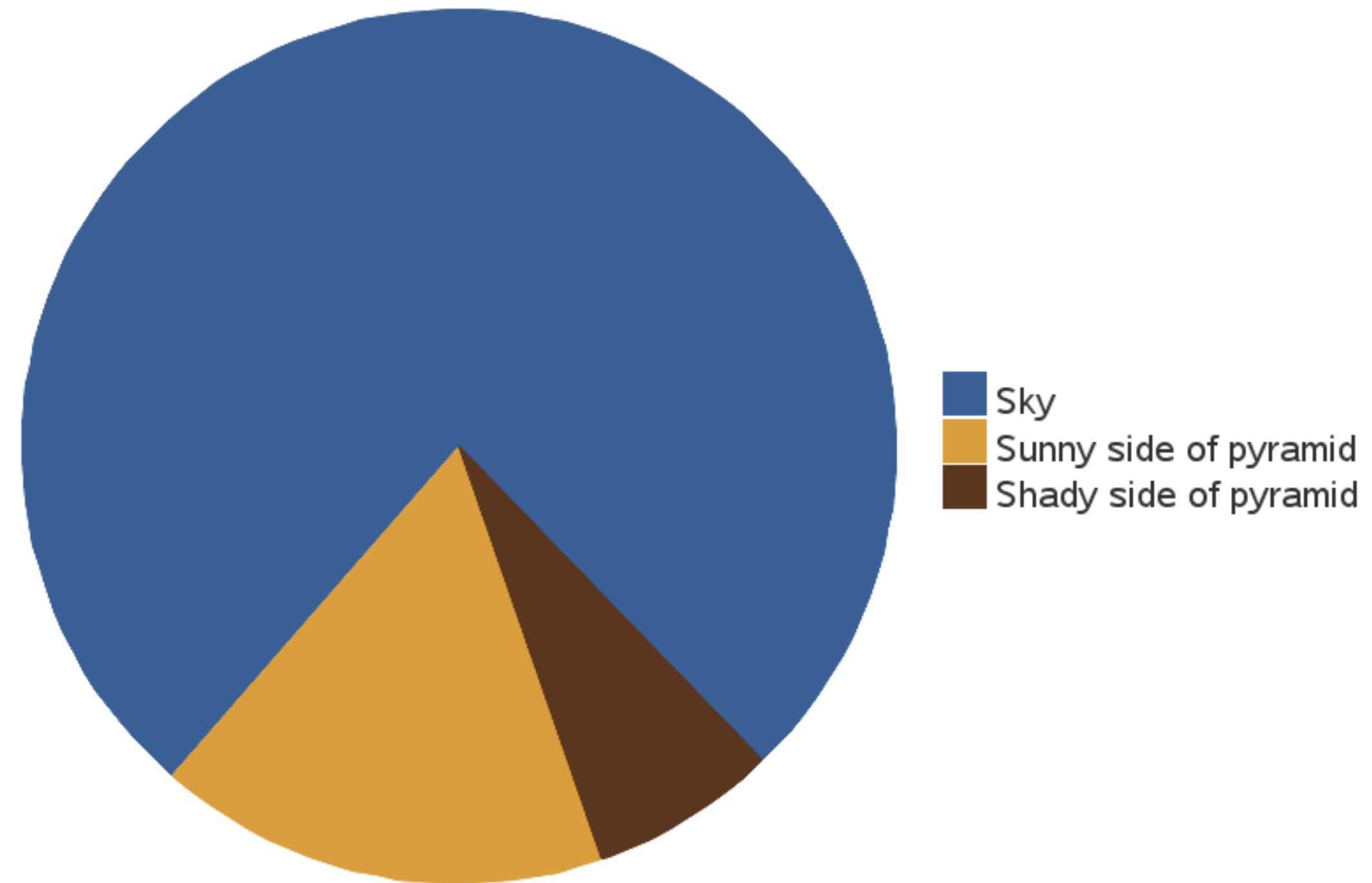
A substitute for pie charts



- Our eyes compare the end points
- Easy to access relative size
- Easy to see how one segment is incrementally larger than other segments
- Labels are right where they're needed

Pie charts

Own a joke category



US financial demographics

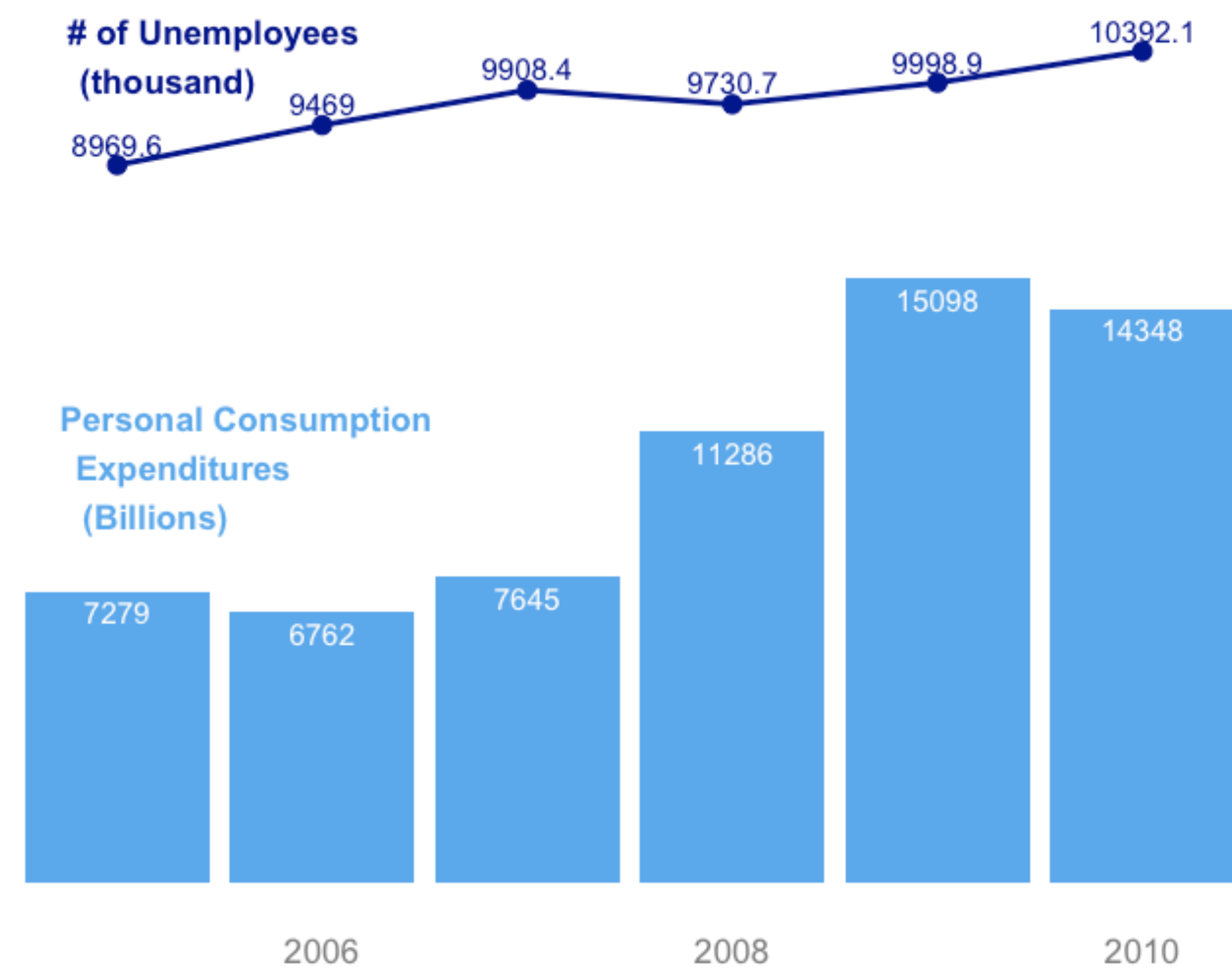
- 1) Are unemployment rates and consumption expenditures negatively correlated?
- 2) US economic time series from Federal reserve bank
- 3) General public

	date	pce	pop	psavert	uempmed	unemploy
1	2015-04-01	12193.8	320402.3	7.6	11.5	8526
2	2015-03-01	12158.3	320230.8	7.4	12.0	8504
3	2015-02-01	12082.4	320074.5	7.9	12.9	8610
4	2015-01-01	12046.0	319928.6	7.7	13.2	8903
5	2014-12-01	12062.0	319746.2	7.6	12.9	8717
6	2014-11-01	12051.4	319564.2	7.3	13.0	9090
7	2014-10-01	12023.0	319353.7	7.2	13.6	8990
8	2014-09-01	11957.4	319125.3	7.4	13.4	9262
9	2014-08-01	11944.3	318893.8	7.2	12.9	9599
10	2014-07-01	11860.5	318662.4	7.5	13.1	9608
11	2014-06-01	11817.0	318464.2	7.4	13.8	9460
12	2014-05-01	11748.4	318269.5	7.4	14.6	9859
13	2014-04-01	11702.6	318089.2	7.4	15.7	9702
14	2014-03-01	11643.0	317917.2	7.4	15.8	10380
15	2014-02-01	11566.2	317753.9	7.3	15.9	10349
16	2014-01-01	11512.5	317593.9	7.1	15.4	10202

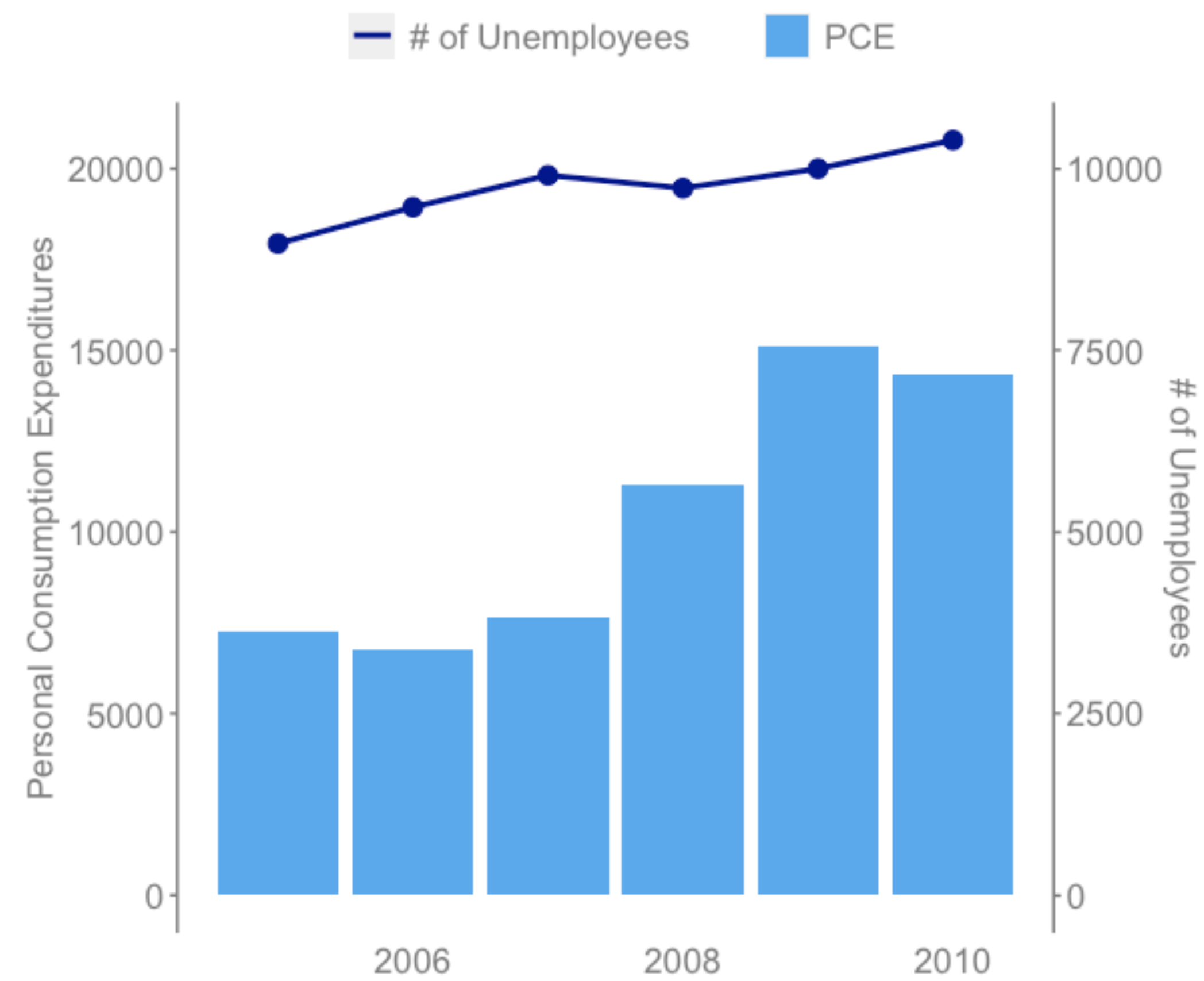
US financial demographics

Avoid displaying a secondary y-axis or a right-hand y-axis

- label the data points that belong on this axis directly
- or
- pull the graphs apart vertically



1



2

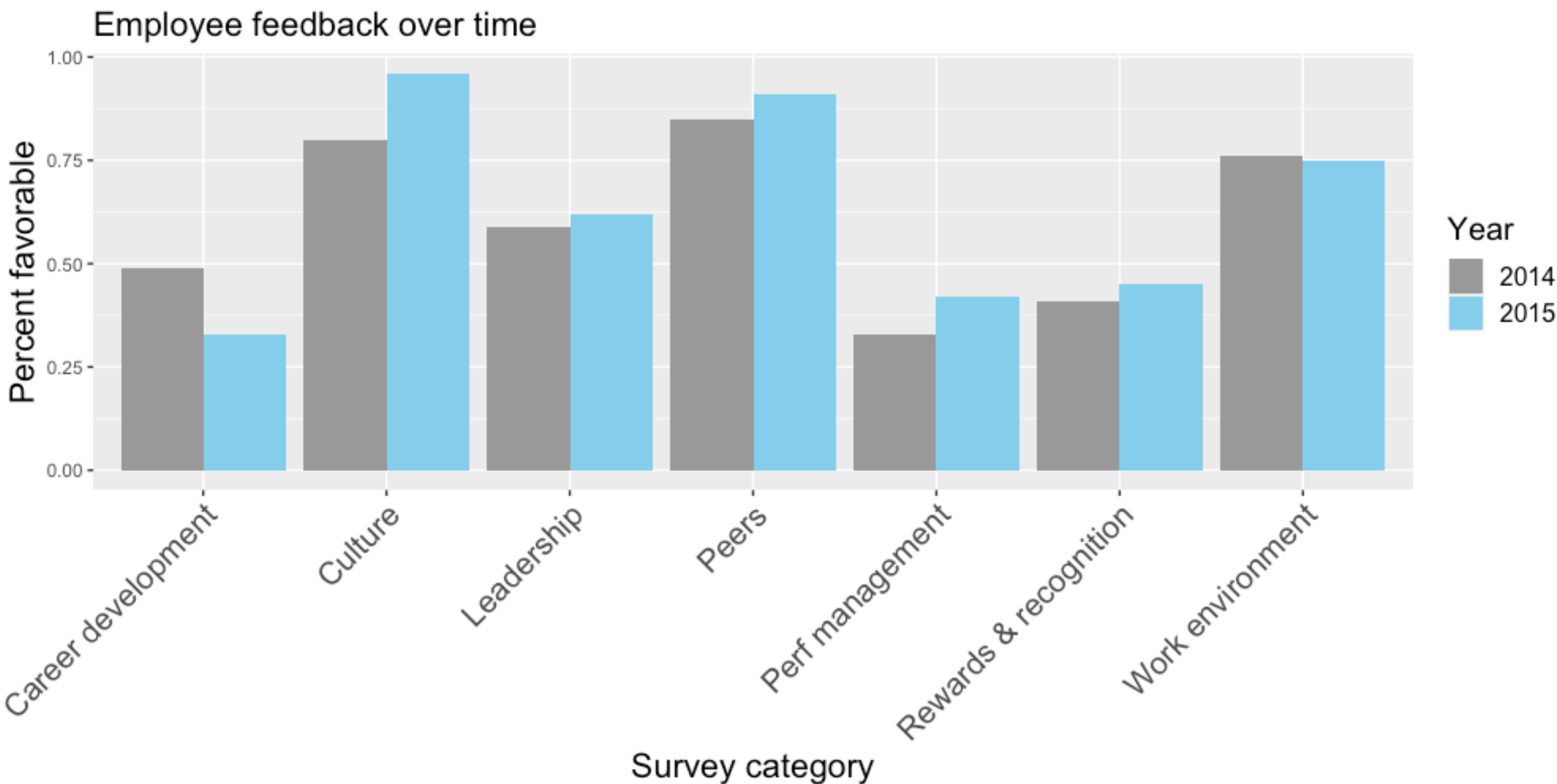
Employee satisfaction

- 1) Explore if our strategies to improve employee satisfaction have made a positive impact
- 2) Results of the satisfaction survey
- 3) HR department

			Satisfaction
	year	category	percent
1	2014	Peers	0.85
2	2014	Culture	0.80
3	2014	Work environment	0.76
4	2014	Leadership	0.59
5	2014	Career development	0.49
6	2014	Rewards & recognition	0.41
7	2014	Perf management	0.33
8	2015	Peers	0.91
9	2015	Culture	0.96
10	2015	Work environment	0.75
11	2015	Leadership	0.62
12	2015	Career development	0.33
13	2015	Rewards & recognition	0.45
14	2015	Perf management	0.42

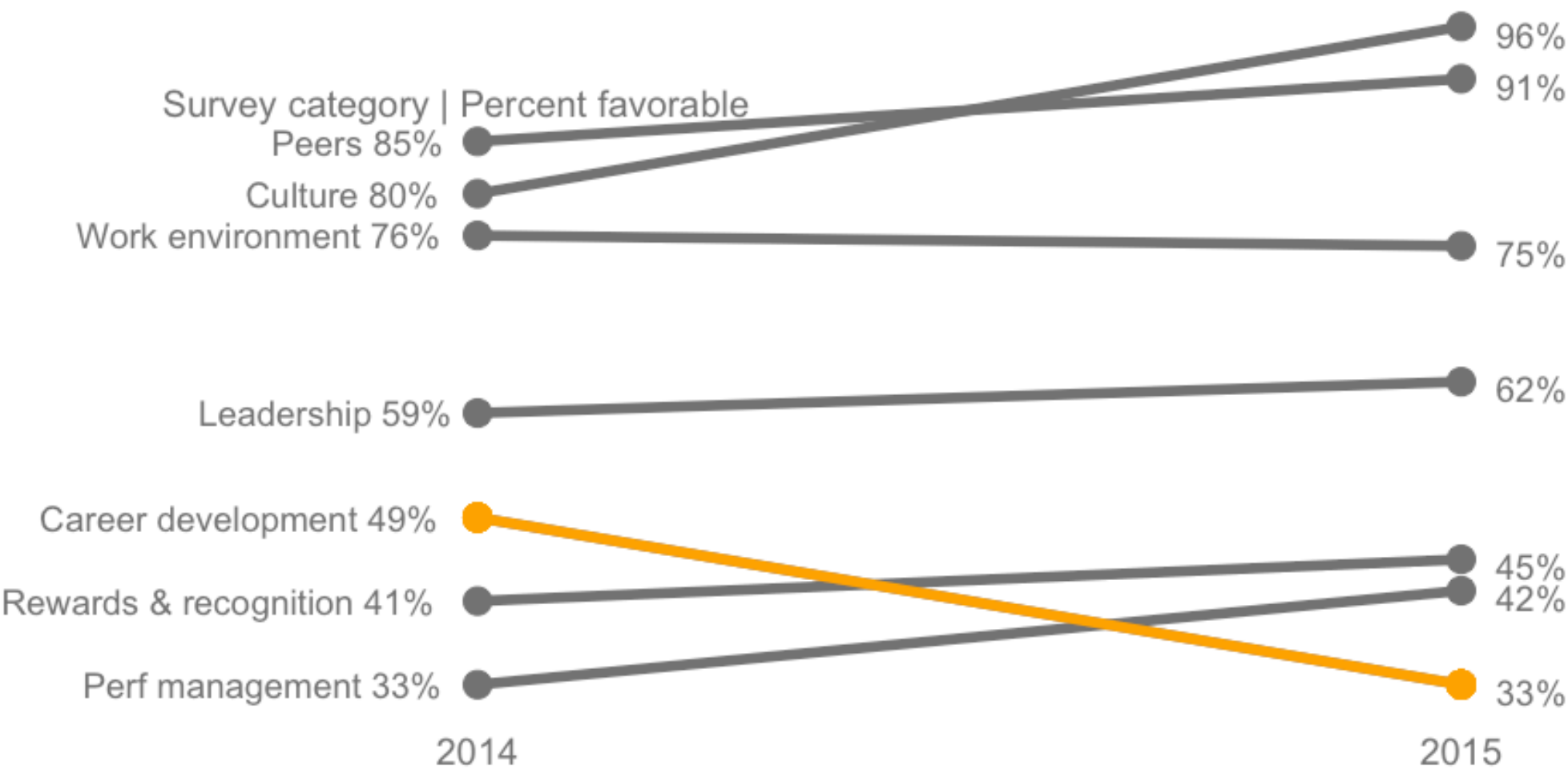
Employee satisfaction

2



1

Employee feedback over time

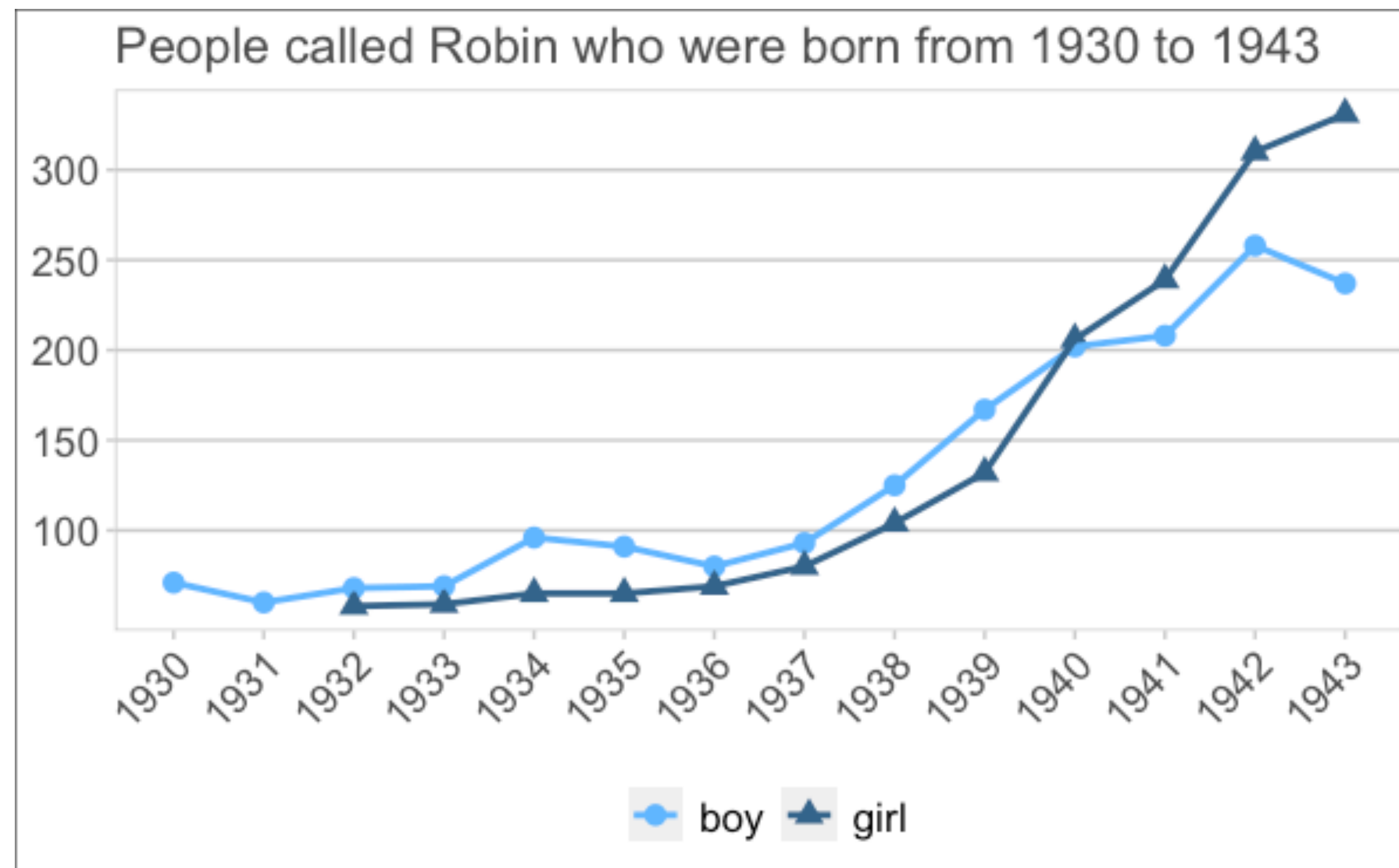


Slopegraph

- When you have two time periods of comparison
- Want to quickly show relative increases and decreases

Plot aesthetics

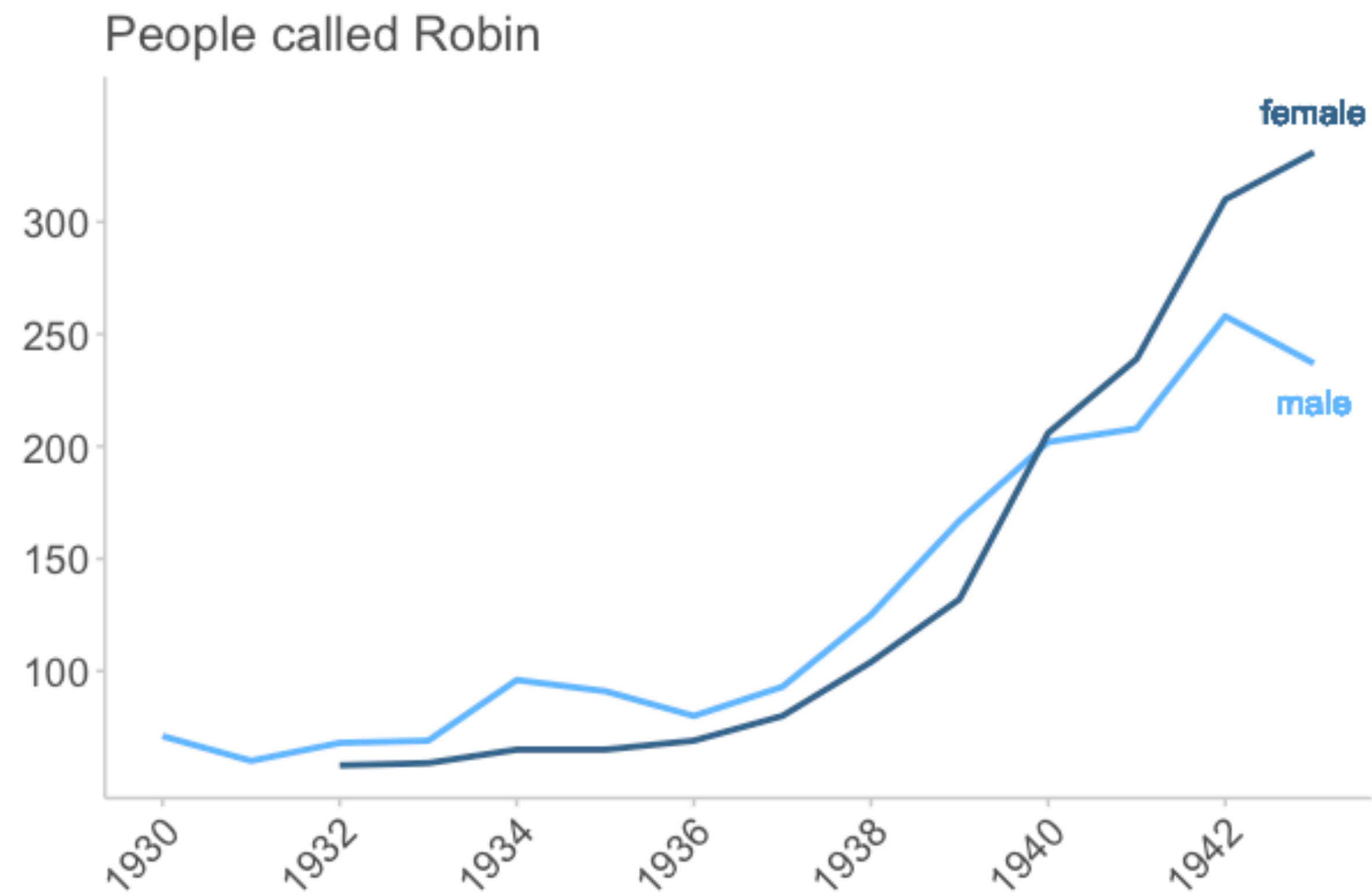
Cluster is your enemy....in general



- Chart border
- Gridlines
- Data markers
- Axis labels
- Redundant title

Plot aesthetics

Cluster is your enemy....in general



- Label data directly
- Leverage consistent color

Questions?