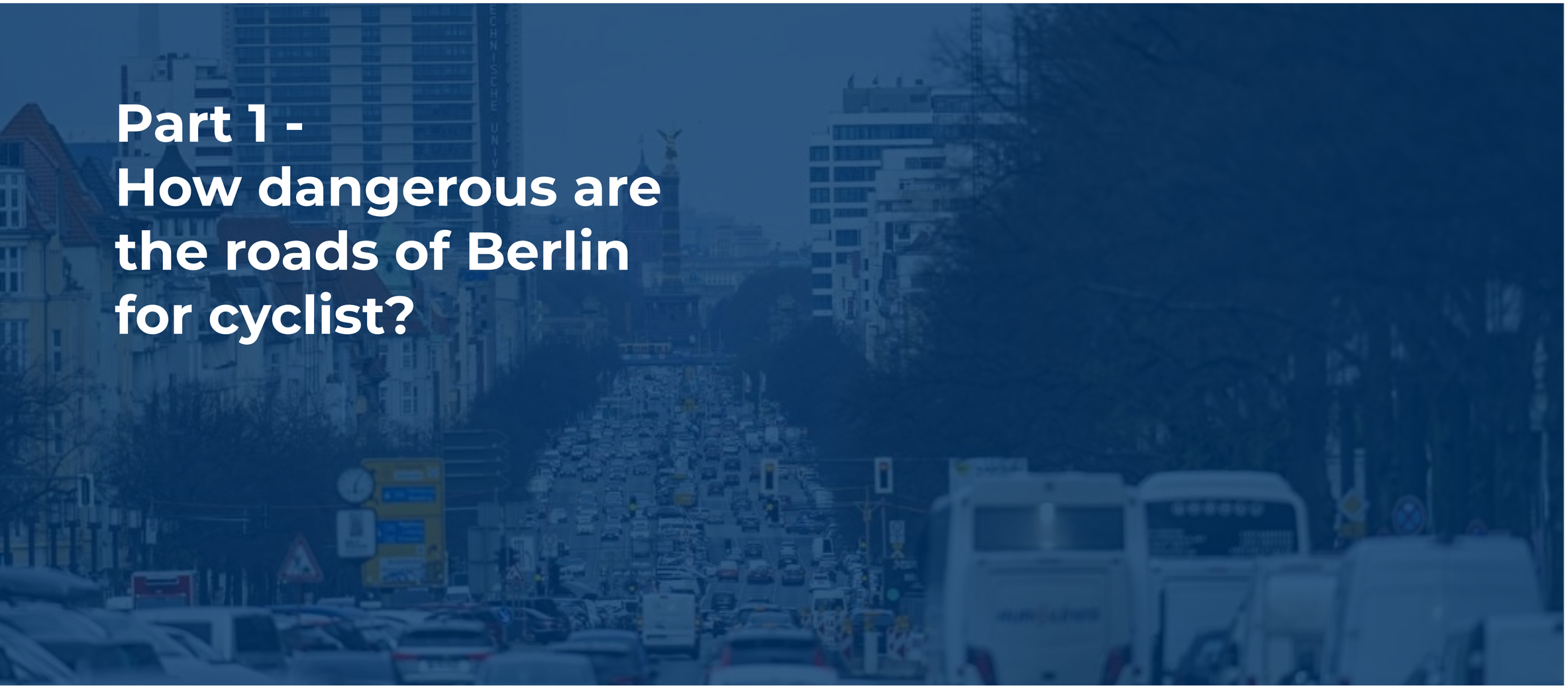
BERLIN IN NUMBERS

PROJECT PITCH DOMENIC SAWATZKI



Part 1 -
How dangerous are the roads of Berlin for cyclist?

questions and **hypothesis**

hypothesis

In districts with a higher population density the percentage of cyclists involved in accidents on the roads is higher.

side questions

Are there roads where do the most accidents occur?

(Top 10 per year)

Are there any pattern between amount of road accidents involving bicycles and the time (date, weekday, month)?

data

berlin road accidents by location in Berlin

- https://daten.berlin.de
- 4 csv files (2018 - 2021)
- approx. 50k entries

berlin population reports

- https://www.statistik-berlin-brandenburg.de
- 8 reports I (2018-2021)

berlin land use, urban structure

- https://daten.berlin.de

mid-term project

time-table

Time slots				Blocks	Data prep **Mo**	Data prep / Investigation **Tue**	Testing / Visualization **Wed**	Visualization / Presenting **Thur**
9:30	**-**	**10:15**	0:45:00	**1**	Get the data	wrangling data	test hypothesis	create visualization (tableau)
10:20	**-**	**11:05**	0:45:00	**2**	Understand tables - data descriptions	wrangling data	test hypothesis	create data story
11:15	**-**	**12:00**	0:45:00	**3**	Cleaning data - road accidents	connect data sets	test hypothesis	create data story
13:00	**-**	**13:45**	0:45:00	**4**	Cleaning data - road accidents	SQL - Database	check side questions	create data story
13:50	**-**	**14:35**	0:45:00	**5**	Cleaning data - population and land use	EDA	check side questions	create presentation
14:45	**-**	**15:30**	0:45:00	**6**	Cleaning data - population and land use	EDA	create visualization (tableau)	create presentation
15:35	**-**	**16:20**	0:45:00	**7**	First investigation - location	EDA	create visualization (tableau)	create presentation
16:30	**-**	**17:15**	0:45:00	**8**	First investigation - location	EDA	create visualization (tableau)	train presentation