



/GROUP 1



Snowstorm East Coast 2010



Julian, Khaldoun, Tina, Yixin



/AGENDA



TOPIC

01

CODING

03

PLOTTING

05



02

PROCESS

04

HYPOTHESES





/ INTRO INTO TOPIC



EST



AST



/ INTRO INTO TOPIC



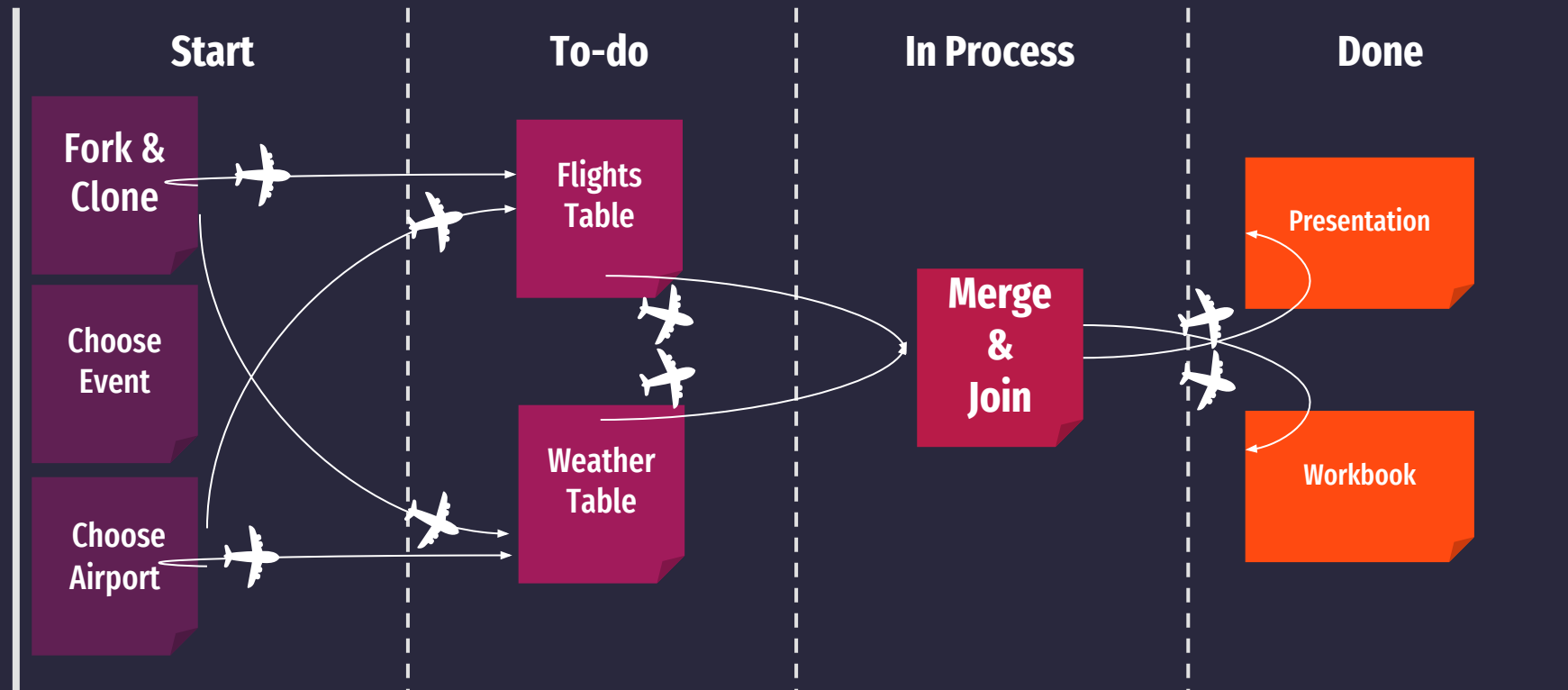
Mon	Tue	Wed	Thu	Fri	Sat	Sun
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

/JANUARY 2011

/DECEMBER 2010						
Mon	Tue	Wed	Thu	Fri	Sat	Sun
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						



/ KANBAN // CODING /// PROCESS





/FILGHTS Table



/GET FLIGHTS DATA

Download and read
csv files

Website:
[https://transtats
.bts.gov](https://transtats.bts.gov)



/DATA CLEANING

Keep and rename
useful columns

Transfer datatype
and clean NULLs



/SQL & JOIN

Connect to SQL
Database + JOIN
on Airport Tables

Read in python



/CLEANING

Data cleaning
again





/WEATHER Table



/GET DAILY STATION DATA

Api: Meteostat (RapidAPI)
Columns: weather_station_id and further weather statistics, e.g. The maximum snow depth in mm



/DATA CLEANING

Replace weather_station_id with airplane_name (for merging)
Keep useful columns
Transfer datatype



/CREATING DATAFRAME

Weather table is ready





/COMBINED Table



HOW:

Left Join

CONDITIONS:

origin_name
(Flights DB) =
airport_name
(Weather DB)
AND
Flight_date
(Flights DB) =
weather_date
(Weather DB)



MERGE:

Flights Table & Weather
Table



DATA CLEANING



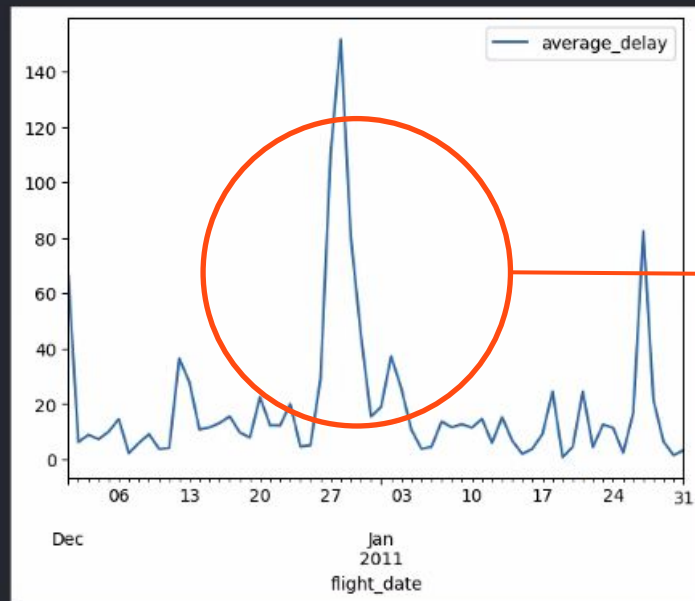


/ HYPOTHESIS (1)

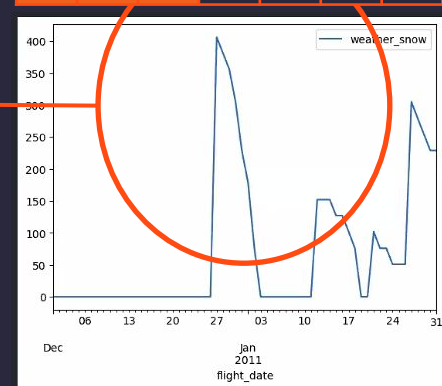


“Is there a discernible correlation between the duration of a snowstorm and the average delay of departure?”

— HYP / GROUP 1



Mon	Tue	Wed	Thu	Fri	Sat	Sun
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		



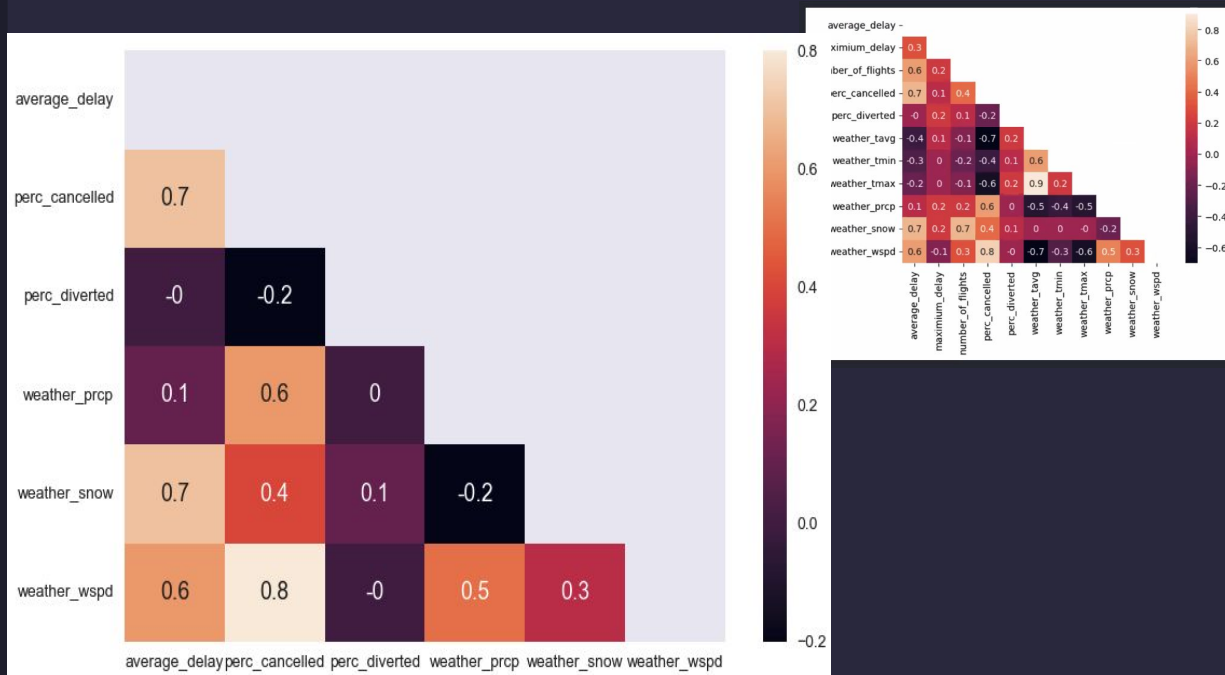


/ HYPOTHESIS (2)



“Among rain, snowfall, and wind strength, which meteorological factor has the strongest correlation with delayed departure time?”

— HYP / GROUP 2



"Focus on success, not distractions."



/ WINDSPEED vs. AVERAGE DELAY

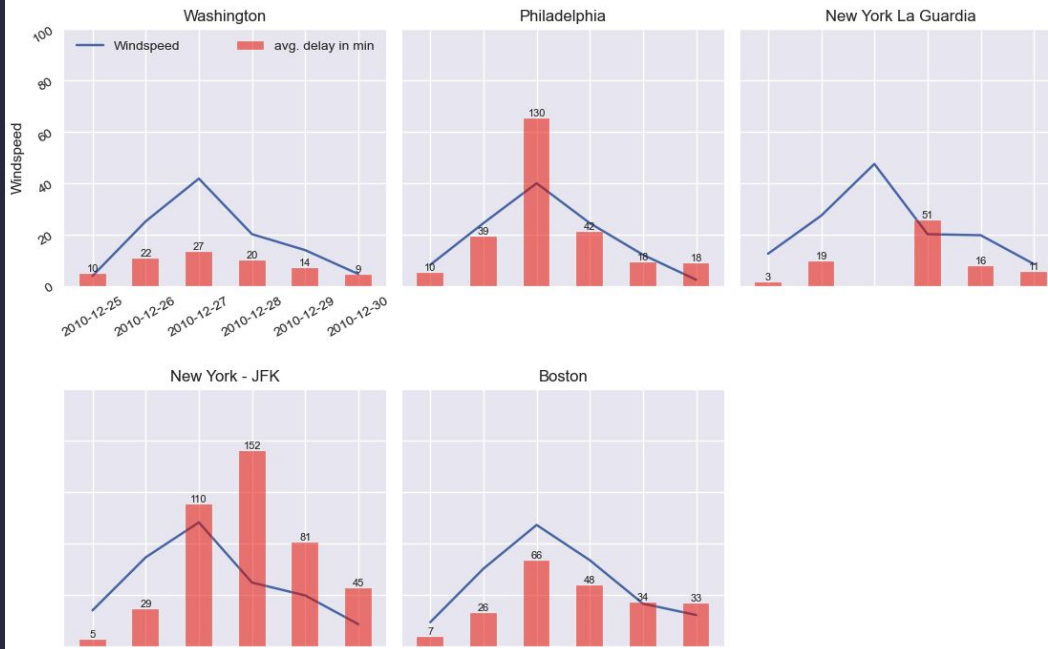


/ CONCLUSION

Correlation does
not imply
causation!



Windspeed vs. Average Delay





/ SNOWFALL vs. AVERAGE DELAY

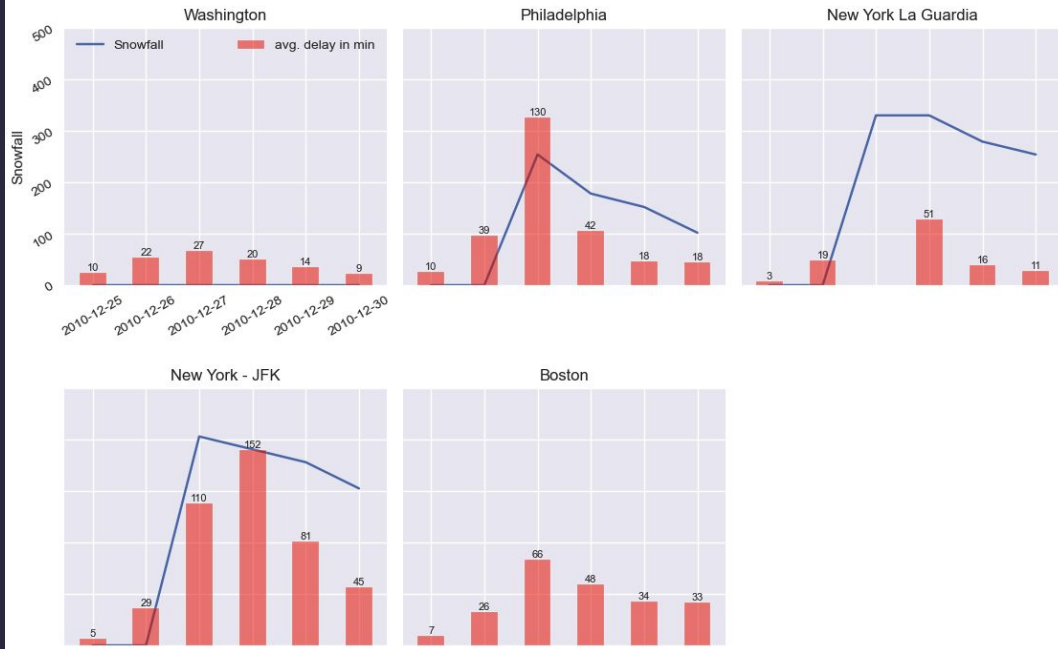


/CONCLUSION

Snowfall is correlated and causally related



Snowfall vs. Average Delay





/ SNOWFALL vs. CANCELLED

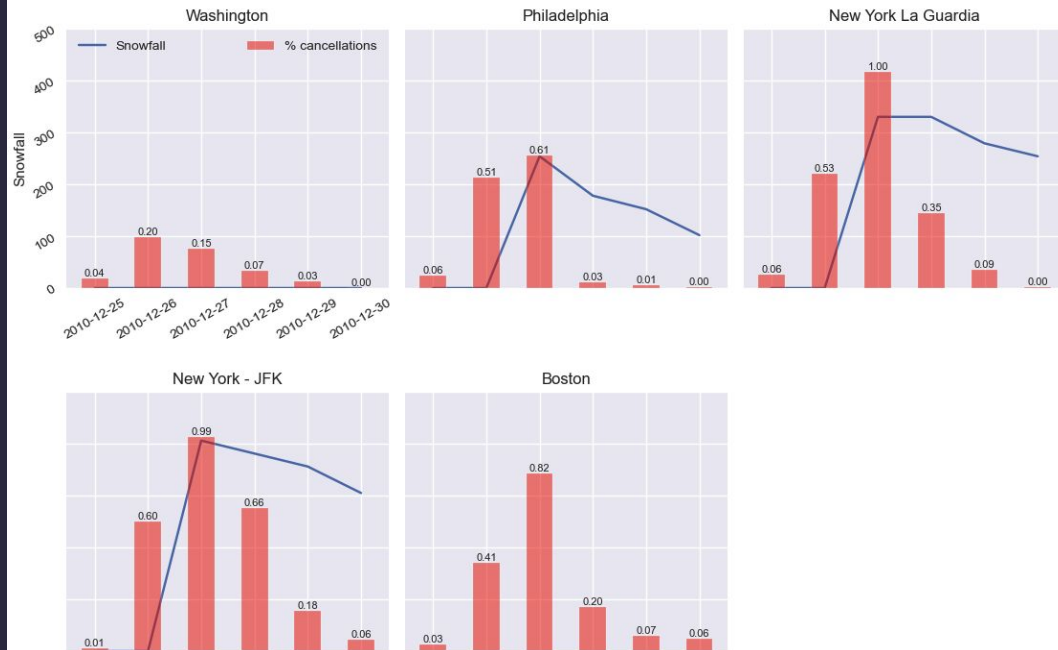


/CONCLUSION

Snowfall is correlated and causally related



Snowfall vs. % Cancelled Flights



/READY TO TAKE OFF?

/DO YOU HAVE ANY QUESTIONS?

/DO YOU HAVE ANY COMMENTS?

/DO YOU HAVE ANY SUGGESTIONS?

/WANT TO SEE OUR NOTEBOOK?



> /NEXT STEPS: EXPAND TIMEFRAME TO 2009 AND COMPARE TO 2010