Turing College: DA

Mod. 1 / Sprint 1 / Proj 1

Deniz Ülker

RQ:

From personal health data,

which hourly interval

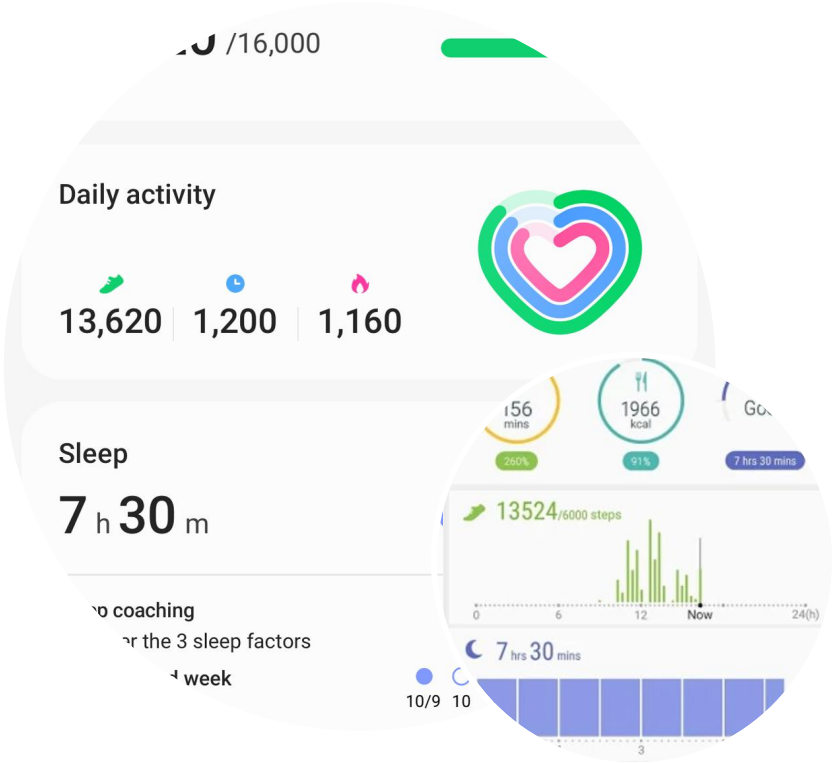
was the most active

during a 5-day period?

Step #0

Exported my own

Samsung Health data



ealth. unt	com.samsung.health. step_count.speed	com.samsung.health. step_count.distance	com.samsu step_cou
	1.3396497	24.57	1.339
	1.0833334	8.76	0.
	1.3888888	9.29	0.
	1.0000001	4.46	0.2
	1.3611112	42.06	2.2
	1.3611112	16.59	0.9

Step #1

Data was messy,

Proceeded with cleaning

> Removed unnecessary fields

> Changed data types i.e.

from ‘General’ to ‘Number’, ‘Date’ etc.

> Reduced decimal places

stepCount	Speed (km/h)	Distance (km)	Calorie (kcal)
32	1.3	24.6	1.3
12	1.1	8.8	0.5
12	1.4	9.3	0.5
6	1.0	4.5	0.3

easurement	Value	Metric
tal steps taken	32404	steps
tal distance traveled	24032.92	km
tal calories burned	1347.25	kcal
ax. Steps	120	steps
g. Steps	54.28	steps
ax. Speed	3.29	km/h
g. Speed	1.23	km/h
ax. Distance	94.76	km
g. Distance	40.26	km
ax. Calorie	6.27	kcal
g. Calorie	2.26	kcal

Step #2

Following variables remained:

- *Step_count*
- *Speed*
- *Distance*
- *Calorie*

> Picked random 5 day interval from entire dataset

> Computed max, avg. & total values

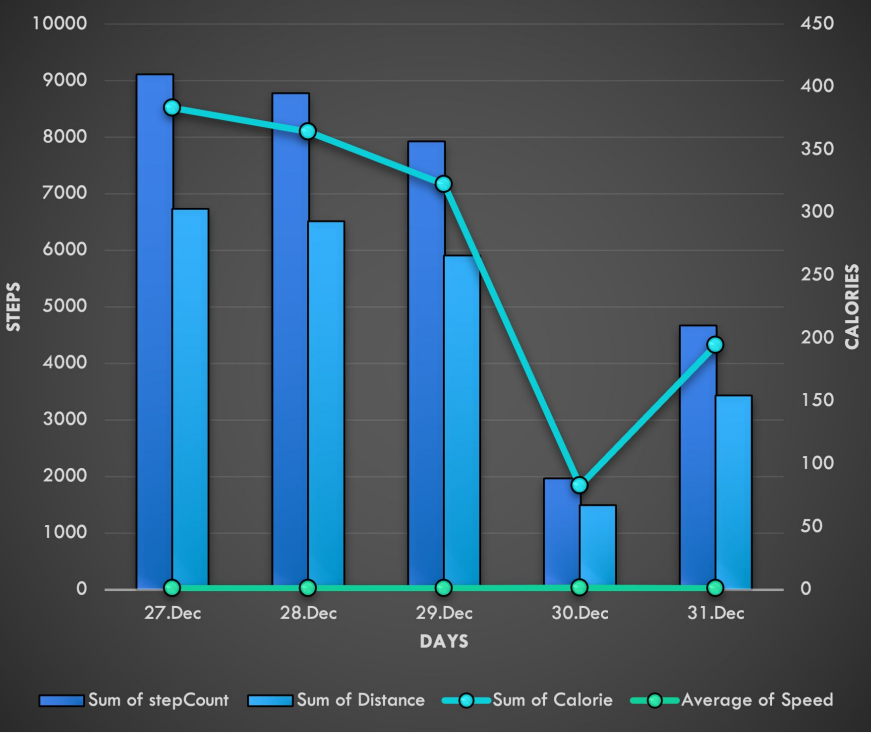
> Next step is to summarize

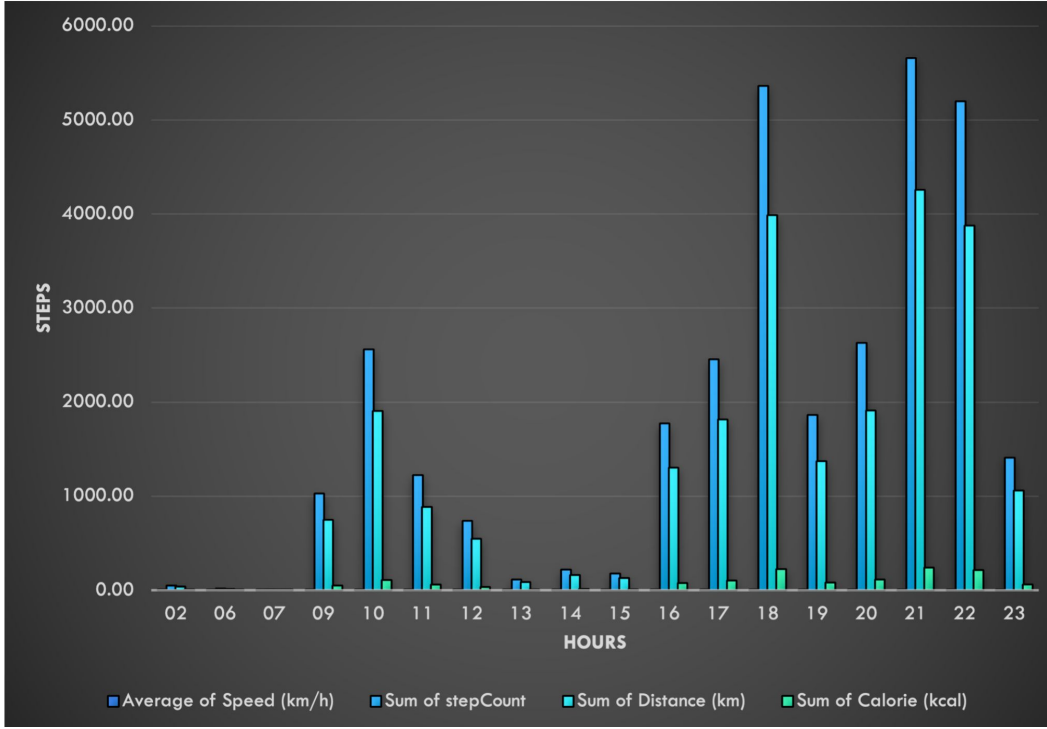
stepCount	Sum of Distance	Sum of Calorie
9098	6721	383
8770	6500	364
7918	5901	322
1958	1485	83
4660	3426	195
32404	**24033**	**1347**

eed (km/h)	Sum of stepCount	Sum of Distance
1.21	44	
1.39	12	
1.00	6	
1.22	1022	
1.27	2557	
1.13	1221	
1.37	734	
1.09	112	
1.17	213	
1.14	171	
1.23	1771	
1.23	2449	
1.20	5358	
1.27	1857	
1.14	2623	
1.30	5655	
1.21	5194	
1.28	1405	
1.23	**32404**	

Step #3

> Summarized data with daily & hourly Pivot Tables.

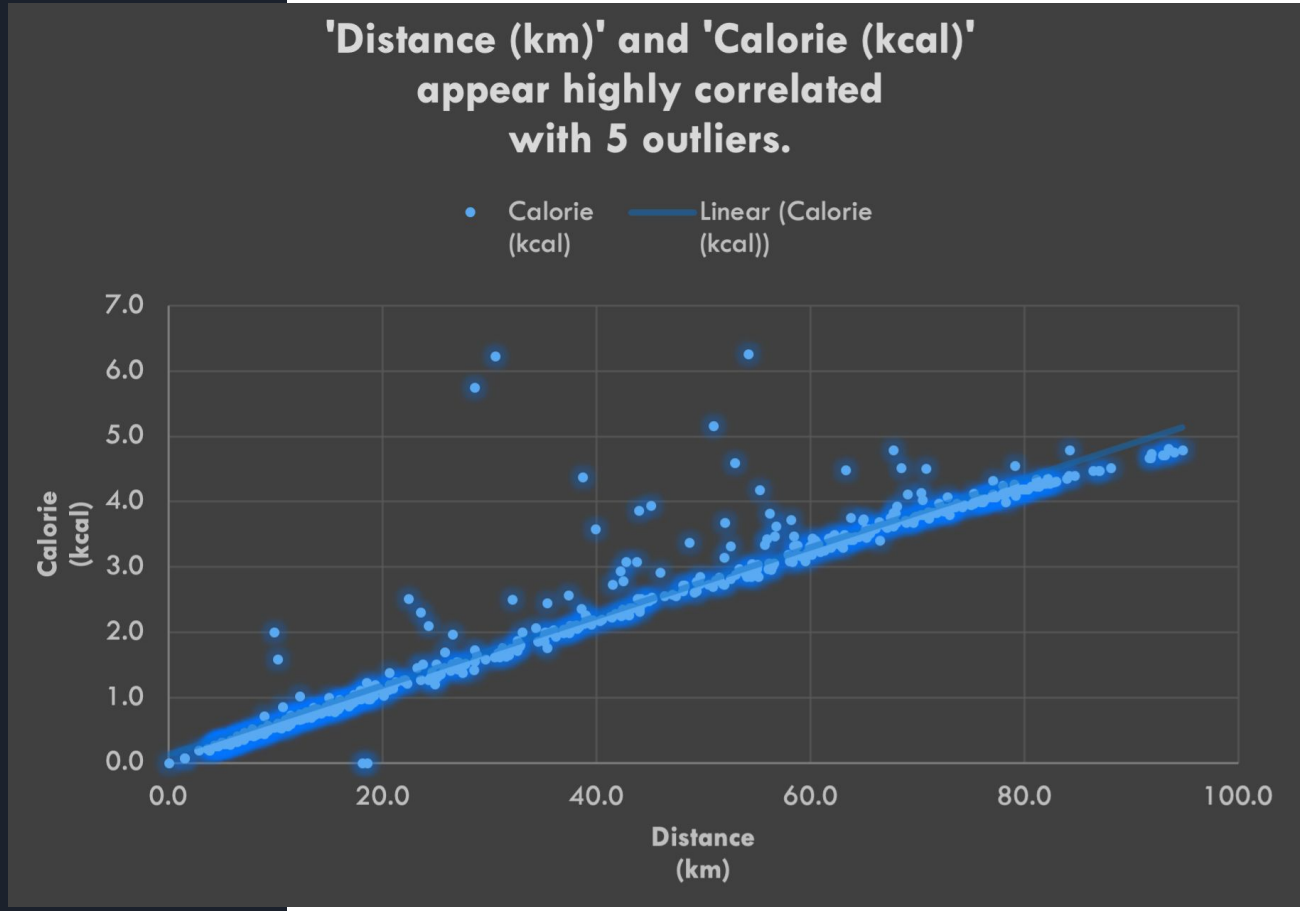




Step #4.1 | Visualisations

Step #4.2 |

Visualisations



Step #5 |

Conclusion

- December 27th was the most active day in terms of Step Count
- Based on aggregation, 21.00 was the most active hour in terms of Step Count & Distance
- As per intuitive deduction, Calorie and Distance are highly correlated
- The same as above is also true for Calorie & Speed (not visualized)