

Date 17/07

Attendees: Alisa, Madina, Amily, Blessing, Hassanat

Discussion matter:

Question we trying to answer?	
	How public transport is affected by the weather in London?
	Public transport in different cities, and factors which affect its reliability
Meetings schedule	everyday stand-ups (10-15 min right after class)
	Sprint retrospective (Sunday 6pm - 7pm)
Next steps	Find data sources for Public Transportat (API, or any that we can scrape)
	Find data for Weather (Api)
	Finalize project topic

Date 18/07

Stand up

Attendees: Alisa, Madina, Amily, Blessing, Hassanat

Question we trying to answer?	
Find data sources for Public Transportat (API, or any that we can scrape)	TFL API is not responding
Find data for Weather (Api)	
Question we trying to answer?	
	is there any correlation between songs lyrics and popularity of the song? If so, what is the formula of the hit?
Next steps	
	Finalize project topic
	Find API for song lyrics (https://www.lyrics.com/lyrics_api.php ,)
	Find API for song popularity (Spotify)

We are going to analyze songs which comply the following criteria: <ul style="list-style-type: none"> - genres (pop music) - date of release (songs released since 2019) - singer (everyone) - language of song (english) 	
responsibilities: <ul style="list-style-type: none"> - spotify API Amily - Lyrics API scraping Alisa, Madina - list of singers scraping(find singer API, website with list of singer) (names, nationality, language, sex, genre) Hassanat, Blessing 	

Date 19/07

Stand up

Attendees: Alisa, Madina, Amily, Blessing

Absent: Hassanat

Day to day updates

Date 20/07 Stand up Attendees: Alisa, Madina, Amily, Blessing Absent: Hassanat		
task	status	reponsible
playlist	<ul style="list-style-type: none"> - found a playlist that is going to be analyzed https://open.spotify.com/playlist/5S8SJdl1BDc0ugpkEvFsIL 	Amily
lyrics	<ul style="list-style-type: none"> - wrote a script that scrapes the lyrics of Lyrics.com API - waiting for the token to use this API 	Alisa, Madina
data analysis	preparing a list of variables (factors) potentially affecting the popularity of the song that we need to gather data	Blessing

	about. (Preparing a spreadsheet template)	
singers info		
Jira	organize the project management platform	Hassanat

Date 21/07 Stand up Attendees: Madina, Amily, Blessing, Hassanat Absent: Alisa		
task	status	reponsible
playlist	<ul style="list-style-type: none"> - some songs have several artists - spotify has several versions of one song, thus different ratings 	Amily
lyrics	<ul style="list-style-type: none"> - token from lyrics.com is not available. decided to go with azlyrics.com - need to update scraping for azlyrics.com 	Alisa, Madina
data analysis	<ul style="list-style-type: none"> - preparing a list of variables (factors) potentially affecting the popularity of the song that we need to gather data about. (Preparing a spreadsheet template) 	Blessing
singers info	<ul style="list-style-type: none"> - find and scrape website with info about artists 	Hassanat, Blessing
JIRA	<ul style="list-style-type: none"> - to explain jira to teammates 	Hassanat

Date 24/07 Sprint retrospective Attendees: Madina, Amily, Blessing, Hassanat, Alisa		
task	status	reponsible
all the useful info from spotify	<ul style="list-style-type: none"> - finished track info - finished artist info - audio feature to be continued 	Amily
lyrics	<ul style="list-style-type: none"> - azlyrics did not work as well, decided to go with https://www.musixmatch.com/ - developed scraping for this website - got our IP addresses banned due to many requests - got just 539 songs lyrics - trying to use proxies. gonna buy 100 IP addresses from https://proxy.webshare.io/ - extract lyrics, and data from lyrics 	Alisa, Madina
data analysis	<ul style="list-style-type: none"> - preparing a list of variables (factors) potentially affecting the popularity of the song that we need to gather data about. (Preparing a spreadsheet template) 	Blessing
singers info	<ul style="list-style-type: none"> - find and scrape website with info about artists 	Hassanat, Blessing
Project management tool	<ul style="list-style-type: none"> - to explain jira to teammates - try notion if jira doesnt work 	everyone

Date 25/07 Stand up Attendees: Amily, Blessing, Hassanat, Alisa Absent: Madina		
task	status	reponsible
create a folder on drive where we can put all the materials relating to this project and those we need to work with		Madina
Work on the Homework 2 together, though at different times and have a final review on Thursday before submitting.		Everyone
Come up with a final table of origin, birthdate and sex by Wed.	<ul style="list-style-type: none"> - We created functions for scraping the singer's infoboxes from Wikipedia - We adjusted the names of the singers in the dataset that returned errors then scraped again - We found a different website to scrape the gender of artistes 	Hassanat and Blessing
Complete the table with all the features could be found on spotify by today.	done	Amily

Date 26/07 Stand up Attendees: Amily, Blessing, Hassanat, Alisa Absent: Madina		
task	status	reponsible
create a folder on drive where we can put all the materials relating to this project and those we need to work with	done	Amily
Work on the Homework 2 together, though at different times and have a final review on Thursday before submitting.		Everyone

Date 27/07 Stand up Attendees: Amily, Blessing, Hassanat, Madina Absent: Alisa		
task	status	reponsible
Trying to get lyrics from AZlyrics using proxies and 50 user-agents. Still get banned every 200 requests Downloaded 600 songs lyrics Alisa will explore selenium		Madina, Alisa
dropping entries where different artists have the same name	Done	Amily
Edit the artist names to be discoverable on wikipedia	Done	Blessing
received dataset from Blessing with cleaned artist name		Blessing and Hassanat

Date 28/07 Stand up Attendees: Amily, Blessing, Hassanat, Alisa, Madina		
task	status	reponsible
create singers' details dataset	Finally completed the dataset containing name, origin and date of birth and gender and it is now ready for cleaning.	Hassanat
Continue downloading lyrics Trying to apply more proxies and multithreading		Madina
Wrote Selenium script that seems to work. Continue to download lyrics.	in progress	Alisa
Assist Blessing and Hassanat with cleaning the singers dataset. (Amily: Yayyy)	in progress	Amily (birth place), Blessing(), Hassanat

Ideally on GitHub you would show that you know how to use version control as a team (e.g., do multiple commits, do PRs with other team members reviewing and maybe asking for corrections; maybe even creating issues). We expect to see that everyone in the team is involved in the project workflow.

Date 31/07 Sprint retrospective Attendees: Madina, Amily, Blessing, Hassanat, Alisa		
task	status	reponsible
running the xlsx file from hassanat to improve algorithm for artist_country extraction		Amily

Date 01/08 Stand-up Attendees: Madina, Amily, Blessing, Hassanat, Alisa		
task	status	reponsible
combine birth_year function from blessing		Amily
-extract lyrics -combine data -start lyrics analysis: # total_char # repititions rate (how many and how much) # unique words (total and relative) # mean word length # chorus repeated # mean length of line # total words	Extracted about 70% of songs from the list. The rest are not available	Alisa, Madina

Date 2/08 Stand-up Attendees: Madina, Amily, Blessing, Hassanat, Alisa		
task	status	reponsible
combine artist_detail file with spotify_api file		Amily
-wrote functions for lyrics analysis: # total_char # repititions rate (how many and how much) # unique words (total and relative) # mean word length # mean length of line # total words	-functions are complete. -next step is to prepare one csv file combined with track_id from Amily's initial csv file	Alisa, Madina

Date 3/08 Stand-up Attendees: Madina, Amily, Blessing, Hassanat, Alisa		
task	status	reponsible
no update		

Date 4/08 Stand-up Attendees: Madina, Amily, Blessing, Hassanat, Alisa		
task	status	reponsible
organize all my code, write comments and explanations, seek for potential better algorithm for everything		Amily
rescrape data such that artist_id will be appended to the scraped data	yet to start	Hassanat

Date 08/08 Stand-up Attendees: Madina, Amily, Blessing, Hassanat, Alisa		
task	status	reponsible
rescrape artists details	finished	Hassanat
concat all the artist details columns		Blessing, Alisa
combine all datasets for our analysis		Amily
create a jupyter notebook and fill-in everything could be done	done	Amily

Date 08/09 Stand-up Attendees: Madina, Amily, Blessing, Hassanat, Alisa		
task	status	reponsible
analysis		Alisa, Amily
upload remaining artist details code, change variable name and file name		Blessing, Hassanat
upload remaining lyrics code, change variable name and file name		Alisa, Madina

Date 08/10 Stand-up Attendees: Madina, Amily, Hassanat, Alisa Absent: Blessing		
task	status	reponsible
analysis		Alisa, Amily
upload remaining artist details code, change variable name and file name		Blessing, Hassanat
upload remaining lyrics code, change variable name and file name		Alisa, Madina

Date 08/11 Stand-up Attendees: Amily, Hassanat, Alisa Absent: Blessing, Madina		
task	status	reponsible
analysis		Alisa, Amily
upload remaining artist details code, change variable name and file name		Blessing, Hassanat

upload remaining lyrics code, change variable name and file name		Alisa, Madina
writing project report		Hassanat

Date 08/13 Stand-up Attendees: Amily, Hassanat, Alisa Absent: Blessing, Madina		
task	status	reponsible
analysis		Alisa, Amily
upload remaining artist details code, change variable name and file name		Blessing, Hassanat
upload remaining lyrics code, change variable name and file name		Alisa, Madina
writing project report		Hassanat, Amily

8-14 finish up & submission