#Proces + kody (Wojtek)

|  |  |
| --- | --- |
| getFBDataFromJSONs.py | Kod getFBuserdata przeszukuje buckety, tworzy listę jsonów i pobiera dane z jsonów do dataframu, odwołuje się do kodów pobierających dane, dane trafiają do df |
| getFBLikesFromJSON.py | pobiera dane z jsona |
| getFBPhotosFromJSON.py | pobiera dane z jsona |
| getFBPostsFromJSON.py | pobiera dane z jsona |
| getFBUserFromJSON.py | pobiera dane z jsona |
| uploadFBDataFromJSONs.py | insert danych |

#

DATAMART

Objaśnienia do nazw zmiennych

1. Sum – sumuje obiekty danego typu
2. Cnt – zlicza obiekty danego typu
3. Total – wszystkie obiekty danego typu w określonej grupie
4. Lastmonth – sposób naliczania = data naliczenie – 30 dni
5. Oldest- najstarsze min(created\_time)
6. Latest- najnowsze max(created\_time)
7. 1:1 zmienna z tabeli docelowej tak jak była
8. Photos, posts, user… - tabela źródłowa z którego pochodzi dana zmienna

|  |  |
| --- | --- |
| Datamartuser | Uwagi: |
| Columns: |  |
| generationdate | Data na dzień naliczenia |
| user\_id | Id |
| sum\_likescnt\_photos | Suma likeów ze zdjęć usera |
| sum\_commentscnt\_photos | Suma comentów ze zdjęć na usera |
| cnt\_image\_name\_photos | ile ze zdjęć usera ma opis |
| users\_total\_photos | Liczba Wszystkie zdjęcia usera |
| oldest\_photos\_age\_days\_photos | Wiek najstarszego zdjęcia na dzień naliczenia w dniacg |
| latest\_photos\_age\_days\_photos | Wiek najnowszego zdjęcia na dzień naliczenia |
| sum\_likescnt\_lastmonth\_photos | Suma likeów ze zdjęć usera (30 dni od daty naliczenia) |
| sum\_commentscnt\_lastmonth\_photos | Suma comentów ze zdjęć na usera (30 dni od daty naliczenia) |
| cnt\_image\_name\_lastmonth\_photos | ile ze zdjęć usera ma opis |
| users\_lastmonth\_photos | Liczba Zdjęć usera wrzucone przez (30 dni od daty naliczenia) |
| age\_group\_13\_17\_user | Klasyfikuje usera do danej grupy wiekowej. Są tylko te 3 grupy wiekowe. |
| age\_group\_18\_20\_user | Klasyfikuje usera do danej grupy wiekowej |
| age\_group\_more\_then\_21\_user | Klasyfikuje usera do danej grupy wiekowej |
| users\_devices\_user | 1:1 |
| users\_currency\_user | 1:1 |
| gender\_male\_1\_user | 1 dla male; 0 dla female |
| test\_group\_user | 1:1 |
| timezone\_user | 1:1 |
| users\_account\_update\_age\_days\_user | Wiek zmian na koncie usera w dniach. Update konta to zmiany wprowadzone w profilu. |
| users\_languages\_cnt\_languages | Ile języków ma user |
| users\_total\_posts | Liczba wszystkich postów |
| total\_comments\_cnt\_posts | Suma wszystkich commentów ze wszystkich postów |
| total\_likes\_cnt\_posts | Suma wszystkich likeów ze wszystkich postów |
| total\_with\_tags\_cnt\_posts | Suma wszystkich with\_tags ze wszystkich postów |
| total\_reactions\_cnt\_posts | Suma wszystkich reakcji ze wszystkich postów |
| oldest\_posts\_age\_days\_posts | Wiek najstarszego postu w dniach na dzień naliczenia |
| latest\_posts\_age\_days\_posts | Wiek najnowszego postu w dniach |
| users\_posts\_cnt\_lastmonth\_posts | Liczba postów usera dla 30dni ostatnich |
| comments\_cnt\_lastmonth\_posts | Liczba commentów usera 30dni ostatnich |
| likes\_cnt\_lastmonth\_posts | Liczba likeów 30dni ostatnich |
| with\_tags\_cnt\_lastmonth\_posts | Liczba With tags dla 30dni ostatnich |
| reactions\_cnt\_lastmonth\_posts | Liczba reakcji dla 30dni ostatnich |
| users\_posts\_reactioncnt\_reactions | Liczba wszystkich reakcji |
| users\_posts\_likecnt\_reactions | Liczba reakcji danego typu |
| users\_posts\_lovecnt\_reactions | Liczba reakcji danego typu |
| users\_posts\_angrycnt\_reactions | Liczba reakcji danego typu |
| users\_posts\_sadcnt\_reactions | Liczba reakcji danego typu |
| users\_posts\_wowcnt\_reactions | Liczba reakcji danego typu |
| users\_posts\_hahacnt\_reactions | Liczba reakcji danego typu |
| users\_posts\_pridecnt\_reactions | Liczba reakcji danego typu |
| users\_posts\_thankfullcnt\_reactions | Liczba reakcji danego typu |
| users\_photos\_reactioncnt\_reactions | Liczba reakcji danego typu |
| users\_photos\_likecnt\_reactions | Liczba reakcji danego typu |
| users\_photos\_lovecnt\_reactions | Liczba reakcji danego typu |
| users\_photos\_angrycnt\_reactions | Liczba reakcji danego typu |
| users\_photos\_sadcnt\_reactions | Liczba reakcji danego typu |
| users\_photos\_wowcnt\_reactions | Liczba reakcji danego typu |
| users\_photos\_hahacnt\_reactions | Liczba reakcji danego typu |
| users\_photos\_pridecnt\_reactions | Liczba reakcji danego typu |
| users\_photos\_thankfullcnt\_reactions | Liczba reakcji danego typu |
| `users\_posts\_reactioncnt\_lastmonth\_reactions`  `users\_posts\_likecnt\_lastmonth\_reactions`  `users\_posts\_lovecnt\_lastmonth\_reactions`  `users\_posts\_angrycnt\_lastmonth\_reactions`  `users\_posts\_sadcnt\_lastmonth\_reactions`  `users\_posts\_wowcnt\_lastmonth\_reactions`  `users\_posts\_hahacnt\_lastmonth\_reactions`  `users\_posts\_pridecnt\_lastmonth\_reactions`  `users\_posts\_thankfullcnt\_lastmonth\_reactions`  `users\_photos\_reactioncnt\_lastmonth\_reactions`  `users\_photos\_likecnt\_lastmonth\_reactions`  `users\_photos\_lovecnt\_lastmonth\_reactions`  `users\_photos\_angrycnt\_lastmonth\_reactions`  `users\_photos\_sadcnt\_lastmonth\_reactions`  `users\_photos\_wowcnt\_lastmonth\_reactions`  `users\_photos\_hahacnt\_lastmonth\_reactions`  `users\_photos\_pridecnt\_lastmonth\_reactions`  `users\_photos\_thankfullcnt\_lastmonth\_reactions` | Liczba reakcji danego typu dla ostatnich 30 dni |
| total\_liked\_by\_user\_likes | Wszystkie obiekty zalikeowane przez usera |
| number\_of\_likes\_in\_popular\_category\_likes | Liczba likeów w Poular\_category –t.j czy częstość występowania danego likea jest większa bądź równa średniej częstości występowania wszystkich likeów |
| total\_categories\_liked\_by\_user\_likes\_category | Liczab wszystkich kategorii likeów polubionych przez usera |
| likes\_in\_popular\_category\_likes\_category | Liczba likeów w Poular\_category –t.j czy częstość występowania danego likea jest większa bądź równa średniej częstości występowania wszystkich likeów |
| user\_all\_tagged\_location\_cnt\_locations | Liczba wszystkich tagów |
| user\_distinct\_location\_cnt\_city\_locations | Liczba wszystkich unikalnych lokalizacji |
| user\_distinct\_location\_cnt\_country\_locations | Liczba wszystkich unikalnych krajów |

Datamart FB Picture

Agregacja na zdjęcie

Sum-sumuje

Cnt-zlicza

|  |  |
| --- | --- |
| Table: DataMartFBPicture |  |
| Columns: |  |
| user\_id | id |
| photo\_id | id |
| from\_id | id |
| generationdate |  |
| created\_time | Data stworzenia |
| image\_name | Nazwa zdjęcia |
| comments\_cnt |  |
| likes\_cnt |  |
| image\_type | Rodzaj zdjęcia (cover,other) |
| av\_Faces | twarze |
| av\_HSV |  |
| av\_Labels |  |
| sam\_sobie | Czy user wrzucił sobie zdjęcie sam czy ktoś mu wrzucił |
| reactioncnt | Liczba Wszystkich reakcji |
| likecnt | Liczba reakcji danego typu |
| lovecnt | Liczba reakcji danego typu |
| angrycnt | Liczba reakcji danego typu |
| sadcnt | Liczba reakcji danego typu |
| wowcnt | Liczba reakcji danego typu |
| hahacnt | Liczba reakcji danego typu |
| pridecnt | Liczba reakcji danego typu |
| thankfullcnt | Liczba reakcji danego typu |
| facescnt | Liczba twarzy na zdjęcie |
| sum\_SAD | Suma danej emocji na zdjęcie |
| sum\_HAPPY | Suma danej emocji na zdjęcie |
| sum\_ANGRY | Suma danej emocji na zdjęcie |
| sum\_confused | Suma danej emocji na zdjęcie |
| sum\_DISGUSTED | Suma danej emocji na zdjęcie |
| sum\_SURPRISED | Suma danej emocji na zdjęcie |
| sum\_CALM | Suma danej emocji na zdjęcie |
| sum\_UNKNOWN | Suma danej emocji na zdjęcie |
| sum\_Sunglasses | Liczba okularów |
| smiletrue | Liczba twarzy uśmiechniętych o prawdopodobieństwie uśmiechu >50% |
| smilefalse | Liczba twarzy uśmiechniętych o prawdopodobieństwie uśmiechu >50% |
| min\_AgeRange\_Low | Wiek najmłodszej osoby na zdjęciu |
| max\_AgeRange\_High | Wiek najstarszej osoby |
| agegap | Przedział wiekowy na zdjęciu |
| avgage | Średni wiek na zdjęciu |
| femalecnt | Liczba kobiet |
| malecnt | Liczba mężczyzn |
| hue | Zmienne z analizy obrazu |
| saturation | Zmienne z analizy obrazu |
| value | Zmienne z analizy obrazu |
| red | Zmienne z analizy obrazu |
| green | Zmienne z analizy obrazu |
| blue | Zmienne z analizy obrazu |