# Paczki – pobieranie danych

|  |  |  |
| --- | --- | --- |
| Nazwa / Link do CRAN | Tematy | Komentarze |
| [eurostat](https://cran.r-project.org/web/packages/eurostat/index.html) | * Ekonomia i finanse * Demografia, warunki życia * Handel zagraniczny * Transport, Energia i środowisko * Społeczeństwo informacyjne | * Dostęp do bazy [Eurostat](https://ec.europa.eu/eurostat/web/main/data/database) * Paczka używa API Eurostat * Dane w formacie długim |
| [OECD](https://cran.r-project.org/web/packages/OECD/OECD.pdf) | * Tematy podobne jak w eurostat | * Dostęp do bazy [OECD](https://stats.oecd.org/) * Zasada działania jak w Eurostat |
| [intsvy](https://cran.r-project.org/web/packages/intsvy/intsvy.pdf) | * Edukacja – wyniki testów PISA, TIMS, PIRLS | * Gotowy wideo [manual](http://www.education.ox.ac.uk/research/r-intsvy-package/) z Oxford * Nie używałem |
| [essurvey](https://cran.r-project.org/web/packages/essurvey/essurvey.pdf) | * Dane z European Social Survey | * Nie używałem |
| [COVID19](https://cran.r-project.org/web/packages/COVID19/index.html) | * Dane pandemiczne – zakażenia, zgodny, liczba łóżek szpitalnych * Indeks uciążliwości restrykcji * Mobilność Google / Apple | * Dane Google Mobility długo się aktualizują – paczka w tle probiera 750MB CSV. |
| [httr](https://cran.r-project.org/web/packages/httr/index.html) | * Paczka, która umożliwia własnoręczne generowanie linków do wszystkich API * Można wykorzystać do pobierania [polskich danych](https://www.gov.pl/web/cyfryzacja/otwarte-dane-dostep-standard-edukacja2) – informacje NFZ, GUS, rejestracje samochodów (CEPIK). | * Otwartedane.gov.pl -wymaga dużej cierpliwości * Informatyzacja pokazuje, co to znaczy państw z dykty. |

# Paczki – social media

|  |  |  |
| --- | --- | --- |
| Nazwa / Link do CRAN | Tematy | Komentarze |
| [rtweet](https://cran.r-project.org/web/packages/rtweet/rtweet.pdf) | * Darmowy dostęp do 3000 tweetów użytkownika oraz statystyk zasięgu. | * Wymaga założenia konta deweloperskiego. |
| [tuber](https://cran.r-project.org/web/packages/tuber/index.html) | * Darmowy dostęp do 3000 zapytań odnośnie statystyk filmów w Youtube | * Wymaga założenia konta deweloperskiego * Brak dostępu do samych filmów (ten można zrobić łatwo w Python) |
| [instaR](https://cran.r-project.org/web/packages/instaR/index.html) | * Treść oraz interakcje z Instagrama. | * Nie używałem |
| [rfacebook](https://cran.r-project.org/web/packages/Rfacebook/Rfacebook.pdf) | * Dane marketingowe z FB | * Wymaga założenia konta deweloperskiego * Nie używałem |

# Dane w Formacie CSV

|  |  |  |
| --- | --- | --- |
| Repozytorium | Przykładowe tematy | Komentarze |
| [Kaggle](https://www.kaggle.com/datasets) | * Baza IMDB z ocenami filmów * Wyniki meczy reprezentacji piłkarskich * Medale olimpijskie | * Realnie duże zbiory danych |
| [Otwarte Dane](https://dane.gov.pl/en) | * Sporo danych rządowych z różnych ośrodków. | * Dobre do ćwiczenia z łączenia kilku plików CSV w jeden zbiór – to jest na każdej rozmowie o pracę w data science. |