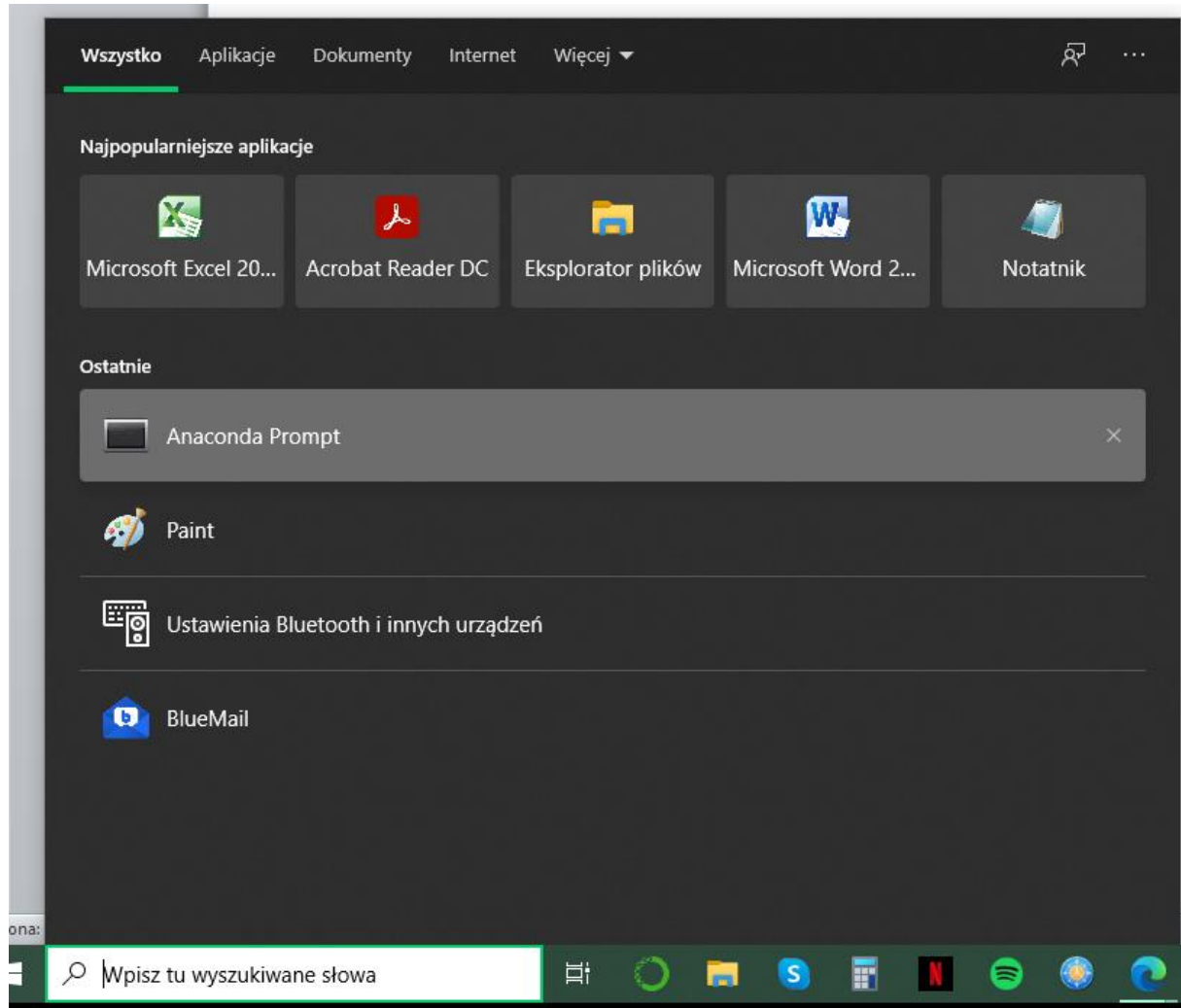


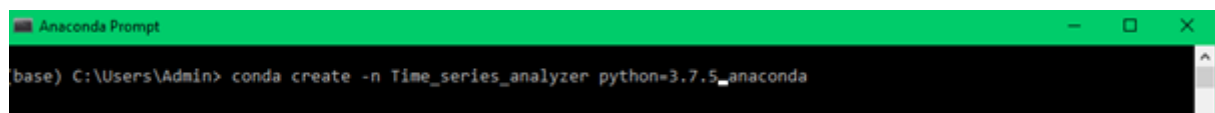
# 1. Instrukcja instalacji programu do analizy szeregów czasowych – instalacja opcjonalna

## 1.1 Wyszukać w programach Anaconda Prompt. Otworzyć.



## 1.2 Stworzyć wirtualne środowisko dla projektu za pomocą komendy

```
conda create -n yourenvname python=x.x anaconda
```



```
conda create -n Time_series_analyzer python=3.7.5 anaconda
```

tworzy środowisko Time\_series\_analyzer w wersji pythona zalecanej do projektu.

Odczekać aż pokaże się ekran podobny po poniższego i zatwierdzić za pomocą **y**

```
Anaconda Prompt - conda create -n Time_series_analyzer python=3.7.5 anaconda

vs2015_runtime pkgs/main/win-64::vs2015_runtime-14.16.27012-hf0eaf9b_3
watchdog pkgs/main/win-64::watchdog-0.10.3-py37_0
wcwidth pkgs/main/noarch::wcwidth-0.2.5-py_0
webencodings pkgs/main/win-64::webencodings-0.5.1-py37_1
werkzeug pkgs/main/noarch::werkzeug-1.0.1-py_0
wheel pkgs/main/noarch::wheel-0.35.1-py_0
widgetsnbextension pkgs/main/win-64::widgetsnbextension-3.5.1-py37_0
win_inet_pton pkgs/main/win-64::win_inet_pton-1.1.0-py37_0
win_unicode_console pkgs/main/win-64::win_unicode_console-0.5-py37_0
wincertstore pkgs/main/win-64::wincertstore-0.2-py37_0
winpty pkgs/main/win-64::winpty-0.4.3-4
wrapit pkgs/main/win-64::wrapit-1.11.2-py37he774522_0
xlrd pkgs/main/win-64::xlrd-1.2.0-py37_0
xlswriter pkgs/main/noarch::xlswriter-1.3.4-py_0
xlwings pkgs/main/win-64::xlwings-0.20.7-py37_0
xlwt pkgs/main/win-64::xlwt-1.3.0-py37_0
xz pkgs/main/win-64::xz-5.2.5-h62dcd97_0
yaml pkgs/main/win-64::yaml-0.2.5-he774522_0
yapf pkgs/main/noarch::yapf-0.30.0-py_0
zeromq pkgs/main/win-64::zeromq-4.3.2-ha925a31_3
zict pkgs/main/noarch::zict-2.0.0-py_0
zipp pkgs/main/noarch::zipp-3.1.0-py_0
zlib pkgs/main/win-64::zlib-1.2.11-h62dcd97_4
zope pkgs/main/win-64::zope-1.0-py37_1
zope.event pkgs/main/win-64::zope.event-4.4-py37_0
zope.interface pkgs/main/win-64::zope.interface-5.1.0-py37he774522_0
zstd pkgs/main/win-64::zstd-1.4.5-h04227a9_0

Proceed ([y]/n)?
```

### 1.3 Aktywować wirtualne środowisko (za każdym razem po odpaleniu anaconda prompt jeśli chce się dogrywać moduły)

Użyć:

conda activate Time\_series\_analyzer

```
Anaconda Prompt

conda activate 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer\' ""
DEBUG menuinst_win32: __init__(196): Menu: name: 'Anaconda${PY_VER} ${PLATFORM}', prefix: 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer', env_name: 'Time_series_analyzer', mode: 'user', used_mode: 'user'
DEBUG menuinst_win32:create(320): Shortcut cmd is C:\Users\Admin\Anaconda3\python.exe, args are ['C:\Users\Admin\Anaconda3\cwp.py', 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer', 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer\python.exe', 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer\Scripts\jupyter-notebook-script.py', "%USERPROFILE%\"]
DEBUG menuinst_win32: __init__(196): Menu: name: 'Anaconda${PY_VER} ${PLATFORM}', prefix: 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer', env_name: 'Time_series_analyzer', mode: 'user', used_mode: 'user'
DEBUG menuinst_win32:create(320): Shortcut cmd is C:\Users\Admin\Anaconda3\pythonw.exe, args are ['C:\Users\Admin\Anaconda3\cwp.py', 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer', 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer\pythonw.exe', 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer\Scripts\spyder-script.py']
DEBUG menuinst_win32:create(320): Shortcut cmd is C:\Users\Admin\Anaconda3\python.exe, args are ['C:\Users\Admin\Anaconda3\cwp.py', 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer', 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer\python.exe', 'C:\Users\Admin\Anaconda3\envs\Time_series_analyzer\Scripts\spyder-script.py', '--reset']
done
#
# To activate this environment, use
#
#     $ conda activate Time_series_analyzer
#
# To deactivate an active environment, use
#
#     $ conda deactivate

(base) C:\Users\Admin>conda activate Time_series_analyzer

(Time_series_analyzer) C:\Users\Admin>
```

Po zmianie środowiska figuruje ono w lewym dolnym rogu w nawiasie ()

## 1.4 Za pomocą komendy **cd** zmienić lokalizację na folder z plikami do programu do analizy danych

W tym przypadku folder ten nazwano **analyzer\_pliki** i umieszczono go na pulpicie.

```
(Time_series_analyzer) C:\Users\Admin>cd Desktop\analyzer_pliki  
(Time_series_analyzer) C:\Users\Admin\Desktop\analyzer_pliki>
```

## 1.5 Po przejściu do folderu użyć komendy **pip3 install -r requirements.txt**

Należy upewnić się, że folder zawiera wspomniany plik. Po użyciu nastąpi automatyczna instalacja modułów wymienionych w pliku.

Może pojawić się komunikat:

```
ERROR: After October 2020 you may experience errors when installing or updating packages. This is because pip will change the way that it resolves dependency conflicts.  
  
We recommend you use --use-feature=2020-resolver to test your packages with the new resolver before it becomes the default.  
  
notebook 6.1.1 requires jupyter-core>=4.6.1, but you'll have jupyter-core 4.4.0 which is incompatible.  
nbclient 0.5.0 requires nbformat>=5.0, but you'll have nbformat 4.4.0 which is incompatible.  
matplotlib 3.3.1 requires certifi>=2020.06.20, but you'll have certifi 2018.8.24 which is incompatible.  
jupyterlab-server 1.2.0 requires jsonschema>=3.0.1, but you'll have jsonschema 2.6.0 which is incompatible.  
jupyter-client 6.1.6 requires jupyter-core>=4.6.0, but you'll have jupyter-core 4.4.0 which is incompatible.  
astroid 2.4.2 requires six>=1.12, but you'll have six 1.11.0 which is incompatible.  
Successfully installed Brotli-1.0.7 Flask-1.0.2 Flask-Compress-1.4.0 Jinja2-2.10 Werkzeug-0.14.1 certifi-2018.8.24 click-6.7  
dash-1.11.0 dash-core-components-1.9.1 dash-daq-0.5.0 dash-html-components-1.0.3 dash-renderer-1.4.0 dash-table-4.6.2 dash-table-experiments-0.6.0 decorator-4.3.0 gunicorn-19.9.0 idna-2.7 itsdangerous-0.24 joblib-0.14.1 jsonschema-2.6.0 jupyter-core-4.4.0 nbformat-4.4.0 numpy-1.19.0 pandas-0.23.4 plotly-4.6.0 plotly-express-0.4.1 python-dateutil-2.7.3 pytz-2018.5 requests-2.19.1 retrying-1.3.3 scikit-learn-0.22.2.post1 scipy-1.4.1 six-1.11.0 traitlets-4.3.2 urllib3-1.23
```

W takim przypadku wykonać:

```
pip install -r requirements.txt --use-feature=2020-resolver
```

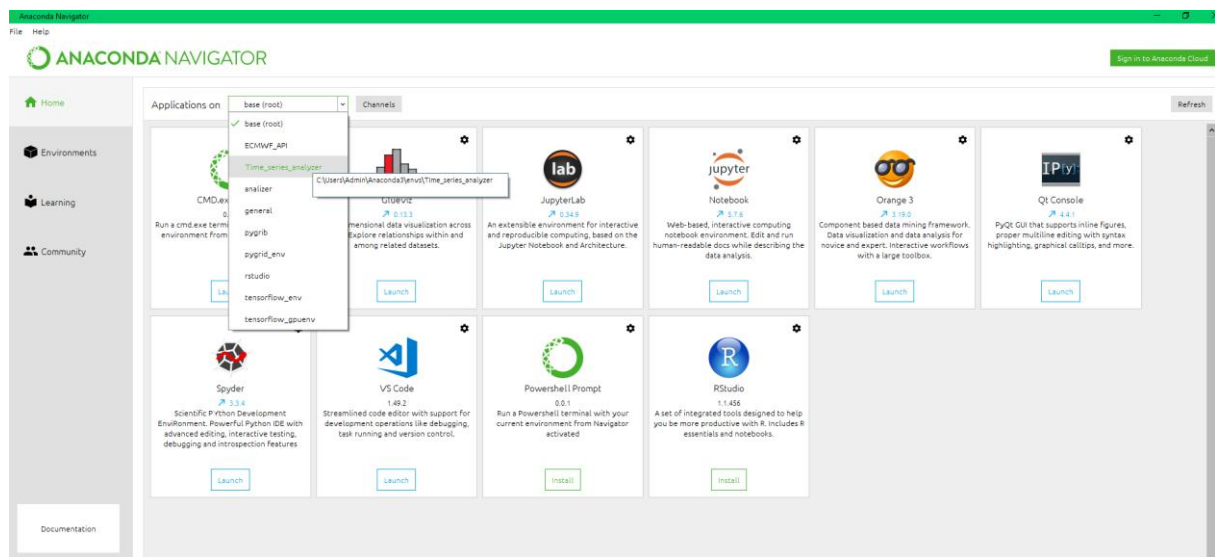
## 1.6 Użyć komendy **python app.py**

Powinien pokazać się tekst podobny do poniższego

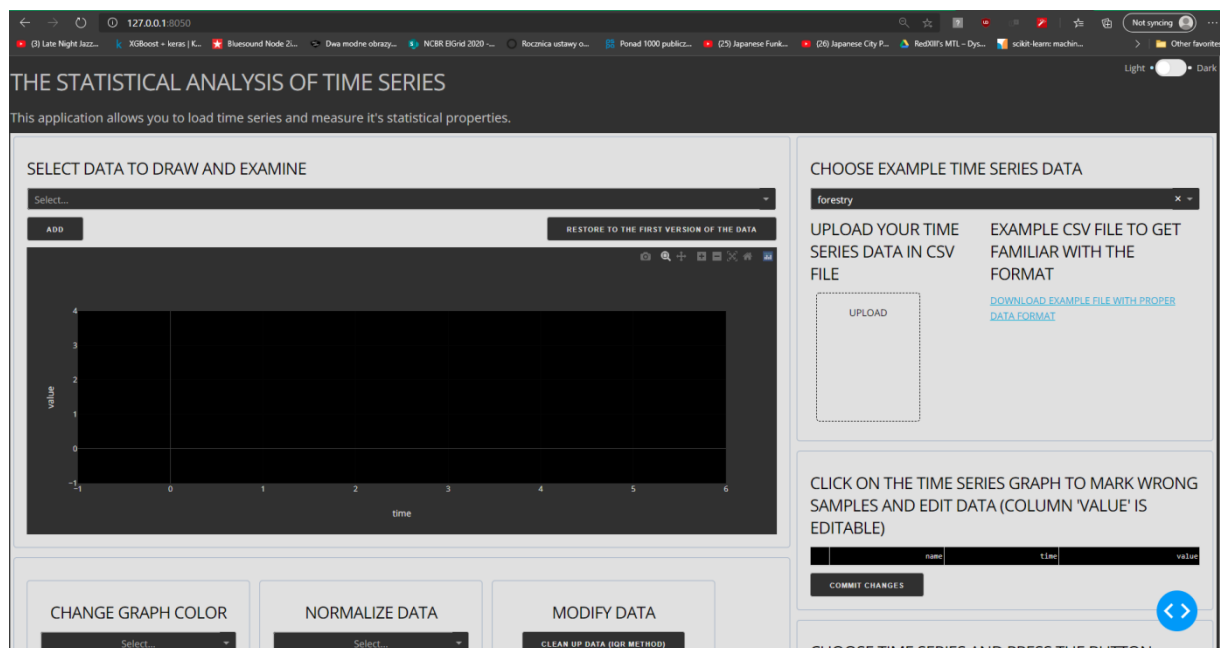
```
(Time_series_analyzer) C:\Users\Admin\Desktop\analyzer_pliki>python app.py  
Running on http://127.0.0.1:8050/  
Debugger PIN: 373-394-491  
* Serving Flask app "app" (lazy loading)  
* Environment: production  
  WARNING: Do not use the development server in a production environment.  
  Use a production WSGI server instead.
```

Otwarcie aplikacji odbywa się przez skopiowanie i wklejenie url do przeglądarki internetowej. W przypadku problemów sugerowane są testy z różnymi przeglądarkami.

->Uruchomić Cmd.exe prompt-> wykonać kroki 4 i 6



Po otwarciu powinno się pokazać:



## 2. Program ćwiczenia

### 0. Uwagi wstępne

Link do wersji webowej programu: [Dash \(time-series-analysis.herokuapp.com\)](https://dash.time-series-analysis.herokuapp.com)

Stosowany w ćwiczeniu program służy zapoznaniu z operacjami istotnymi w analizie i obróbce danych dot. Szeregów czasowych, tj.:

- Normalizacją danych
- Oczyszczaniem oraz wygładzaniem danych
- Zmianą granulacji czasowej danych/eliminacją pomiarów co n-tą próbkę
- Filtracją danych wg. Parametrów czasowych
- Generacją statystyk opisowych, w tym miar pozycyjnych, miar zmienności, asymetrii, rozpiętości, czy koncentracji szeregów czasowych
- Generacją wykresów wizualizujących zależności w danych, w tym wykresy korelacji, wykresy pudełkowe i skrzypcowe
- Szacowaniem parametrów mogących mieć znaczenie przy stosowaniu metod statystycznych prognoz, tj, szacowanie typów rozkładu i określaniu stacjonarności

Program ma zastosowanie wyłącznie dydaktyczne i nie jest rozwijany komercyjnie, stąd też posiada pewne ograniczenia co do wydajności. Celem narzędzia jest pokazanie, jakie, możliwie proste narzędzie można wykonać przy użyciu obecnie istniejących wysokopoziomowych języków programowania i jakie takie technologie mają ograniczenia. Instalacja programu jest jednak dobrowolna. W ramach zajęć pokazane zostaną metody osiągnięcia pożądaných wykresów i miar statystycznych oraz podstawowe metody interakcji z danymi. W tym celu zastosowany będzie język Python. Laboratorium dostosowane jest do posługiwania się tym właśnie językiem wraz z zastosowaniem narzędzi Jupyter notebook lub Spyder. Sposób instalacji Pythona wraz z Jupyter notebookiem został opisany w pliku ĆW 6,7 keras instrukcja zaktualiz.pdf (vide załączniki do kursu).

W przypadku lepszej znajomości innych narzędzi możliwe jest zastosowanie przez wykonujących ćwiczenie dowolnego innego narzędzia umożliwiającego wykonanie programu ćwiczenia.

Każda grupa otrzymuje komplet plików z danymi, tj. pliki:

- signals
- estimated\_inpatient\_covid,
- cov\_deaths,
- gdp-per-capita-worldbank

**Uwaga! Poza plikami signals, dane są w formie surowej – mają więc postać i zawartość w jakiej udostępniła je organizacja je wystawiająca do informacji publicznej. Sugerowany jest krótki przegląd danych w celu wyłapania ewentualnych niespójności itp.**

## **WYKONANIE ĆWICZENIA**

1. Każda grupa wybiera 1 sygnał z pliku signals, trzy kraje z pliku cov\_deaths (jeśli rozbić na prowincje, należy dodać szeregi czasowe) i te same kraje z pliku gdp-per-capita-worldbank, oraz 3 stany w USA z pliku 'estimated\_inpatient\_covid'
2. Następuje ogląd danych i próba dostosowania szeregów czasowych do postaci akceptowanej przez program
3. Dla szeregów należy wykonać oczyszczanie z obserwacji odstających a następnie:
  - a) wizualizację przebiegu czasowego
  - b) histogram
  - c) macierz korelacji szeregów z innymi szeregami wybranymi (nie dla pliku signals)
  - c) określić właściwości statystyczne szeregów czasowych znajdujące się w statystykach obliczanych przez program (średnia, percentyle, etc.)
  - d) porównanie szeregów występujących w obrębie danego pliku
  - e) wnioski który kraj radzi sobie najlepiej i w którym student przewiduje problemy

W przypadku stwierdzenia, że dla przedstawionego problemu, dany typ wykresu nie ma racji zastosowania, można nie generować danego wykresu, co jednak należy odpowiednio uzasadnić. Nacisk w sprawozdaniu należy położyć na część wnioskowania.