**Intel® AI DLA MŁODZIEŻY 2019**

PROGRAM PILOTAŻOWY

# PODRĘCZNIK MODERATORA

**Moduł 16 - Doświadczenie [Dane]**

**(Import i przetwarzanie danych)**

CAŁKOWITY CZAS TRWANIA SESJI: 480 MINUT

LICZBA INSTRUKTORÓW: 2

LICZBA UCZESTNIKÓW: 40

**UWAGA:** ŻADNA OSOBA SPOZA ORGANIZACJI INTEL CORPORATION NIE MOŻE KOPIOWAĆ, ROZPROWADZAĆ, REPRODUKOWAĆ ANI W ŻADEN INNY SPOSÓB UDOSTĘPNIAĆ INFORMACJI ZAWARTYCH W NINIEJSZYM DOKUMENCIE OSOBOM, INSTYTUCJOM I/LUB ORGANIZACJOM NIEPOWIĄZANYM Z NINIEJSZYM PROGRAMEM BEZ ZGODY WYŻEJ WYMIENIONEJ ORGANIZACJI.

NIEPRZESTRZEGANIE POWYŻSZYCH ZASAD BĘDZIE SKUTKOWAĆ ZAKOŃCZENIEM WSPÓŁPRACY PRZEZ WYŻEJ WYMIENIONĄ ORGANIZACJĘ.

NINIEJSZY DOKUMENT MUSI **ZOSTAĆ ZWRÓCONY** UPOWAŻNIONEMU PERSONELOWI FIRMY INTEL CORPORATION PO ZAKOŃCZENIU SESJI.

## **Podręcznik moderatora**

|  |  |
| --- | --- |
| **Temat lekcji:** Moduł 16 - Doświadczenie [Dane]  (Import i przetwarzanie danych) | **Tryb:** Warsztaty w tempie własnym uczestników, ćwiczenia oparte na projektach |
| **Podsumowanie:**  Ten moduł obejmuje umiejętności pozyskiwania i przetwarzania danych. Uczestnicy będą pobierać surowe dane i będą je przetwarzać z uwzględnieniem danych brakujących, wartości odstających i błędów. | |
| **Cel edukacyjny:**   1. Ćwiczenie pozyskiwania danych z użyciem Python 2. Ćwiczenie technik eksploracji danych z użyciem Python 3. Ćwiczenie technik przetwarzania danych z użyciem Python | |
| **Rezultaty edukacyjne:**   1. Umiejętność importowania danych z użyciem Python 2. Umiejętność zautomatyzowania pobierania danych 3. Umiejętność korzystania z percentyli 4. Znajomość wykresów pudełkowych i histogramów 5. Znajomość różnicy między błędami a wartościami odstającymi 6. Umiejętność postępowania z danymi błędnymi i brakującymi | |
| **Wymagania wstępne:**   1. Podstawowa znajomość programowania 2. Podstawy znajomość zagadnień statystyki | |
| **Kluczowe pojęcia:**   1. Gromadzenie danych 2. Przetwarzanie danych | |
| **Kluczowe umiejętności:**   1. Dokładne zrozumienie procesu przetwarzania danych brakujących i błędnych | |
| **Wykorzystywane materiały:**   1. [Slajdy] Module 16 - Doświadczenie [Dane] (Import i przetwarzanie danych) 2. [Arkusz] Module 16 - Doświadczenie [Dane] (Import i przetwarzanie danych) 3. (1) [Notatnik Jupyter - Uczestnicy] Moduł 16 - Pozyskiwanie danych 4. (2) [Notatnik Jupyter - Uczestnicy] Moduł 16 - Podstawy przetwarzania danych i wizualizacji 5. (3) [Notatnik Jupyter - Uczestnicy] Moduł 16 - Postępowanie z danymi błędnymi i brakującymi | |
| **Omówione kwestie etyczne/społeczne**   1. Prawa autorskie i prywatność danych | |
| **Zastosowanie w scenariuszach z prawdziwego życia:**   1. Umiejętność gromadzenia, przetwarzania i zrozumienia danych podczas pracy na projektem z zakresu nauki o danych | |

## **1. Konspekt**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Nie** | **Czas (min)** | **Czynność** | **Opis** | **Cel** |
| 1. Przedstawienie uczestników i „przełamanie lodów” | 30 | Dyskusja moderowana | - Uczestnicy będą się komunikować bez używania języka mówionego.  - Język jako istotny sposób przekazywania informacji. | Moderatorzy poznają uczestników, a uczestnicy poznają siebie nawzajem i moderatorów. |
| 2. Powtórka: Gromadzenie i przetwarzanie danych | 10 | Wykład | Powtórka ze sposobów gromadzenia i przetwarzania danych. | Przypomnienie sposobów gromadzenia i przetwarzania danych. |
| 3. Samokształcenie | 60 | Pozyskiwanie danych | - Pobranie i import danych do Python  - Nauka zastosowań biblioteki Pandas | Umiejętność pozyskania i przetwarzania danych za pomocą Python i Pandas |
| 30 | Kodowanie sekwencyjne | - Utworzenie grup w celu automatycznego pobierania, importu i wyświetlania danych.  - Kod musi być zapisany sekwencyjnie (w turach) |
| 60 | Podstawy przetwarzania i wizualizacji danych | - Korzystanie z podstawowych narzędzi statystycznych (percentyle, wykresy pudełkowe, histogramy i wykresy punktowe) w celu uzyskania przeglądu zbioru danych |
| 60 | Postępowanie z danymi błędnymi i brakującymi | - Zrozumienie różnicy między danymi błędnymi a odstającymi  - Zastosowanie różnych metod postępowania z danymi błędnymi i brakującymi |
| 4. Projekt | 40 | Projekt – Część 1 z 2 | Prezentacja kilku przykładów przed rozpoczęciem wyzwania. 10 zespołów 4-osobowych. | Zastosowanie umiejętności zdobytych podczas sesji. |
| 20 | Kontrola w połowie drogi | Kontrola statusu projektów w połowie czasu, doradcy udzielają porad/wskazówek. |
| 40 | Projekt – Część 2 z 2 | Zespoły dopracowują i testują swoje zastosowania. |
| 40 | Prezentacja projektu | Zespoły prezentują swoje prace i dowiadują się, na jakie pomysły wpadli inni. Uczestnicy dzięki temu uświadamiają sobie, że można na kilka różnych sposobów podejść do rozwiązania tego samego problemu. | Uczestnicy prezentują swoje projekty i uczą się od siebie nawzajem. |
| 5. Podsumowanie | 30 | Podsumowanie | Podsumowanie zdobytej wiedzy, podkreślenie przykładów z projektu. | Utrwalenie wiedzy technicznej. |
| 6. Quiz | 30 | Quiz | Zaprezentuj pytania uczestnikom, aby wspólnie spróbowali na nie odpowiedzieć. | Oceń, w jakim stopniu uczestnicy rozumieją to, czego się nauczyli. |
| 7. Refleksja | 30 | Refleksja | - Rozdaj puste kartki każdemu zespołowi.  - Przypnij refleksje uczestników na tablicy w odpowiednim momencie. | Utrwalenie wiedzy i zakończenie sesji. |

## **2. Przygotowanie sesji**

### Przeczytaj podręcznik moderatora i spróbuj odpowiedzieć na pytania w notatniku Jupyter. Uważnie zapoznaj się z materiałami. Poświęć na to odpowiedni czas; jeśli to twój pierwszy raz, może to zająć kilka godzin.

### Sprawdź, czy na sali jest sprawny projektor na potrzeby wyświetlenia instrukcji dotyczących notatnika Jupyter.

### Skopiuj katalog „**[Pliki robocze – Uczestnicy] Moduł 16 - Doświadczenie [Dane] (Import i przetwarzanie danych)**” zawierający notatniki Jupyter na wszystkie laptopy.

### Aby zapewnić szybkie usunięcie błędów, jakie mogą wystąpić podczas instalacji nowych bibliotek w Python, dopilnuj, by uczestnicy zawsze pracowali w środowisku wirtualnym. Zob. „**Instrukcje dotyczące środowisk wirtualnych.docx**”. Zrestartuj nowe środowisko wirtualne, jeśli wystąpią błędy podczas instalacji.

### Czas podany dla każdego ćwiczenia i notatnika jest orientacyjny. Moderatorzy mogą wydłużać lub skracać poszczególne zadania według własnego uznania. Najważniejszym celem jest, aby uczestniczy w pełni zrozumieli wykonywane zadania i ćwiczenia.

### Inne wymagane materiały

### Markery do pytań w dyskusji po ostatnim ćwiczeniu grupowym

### Stoper (wystarczy stoper w telefonie. Posłuży on do mierzenia czasu w trakcie zadań grupowych.

## **3. Przewodnik po czynnościach**

### ***[Slajdy 1–4] Wprowadzenie i przełamanie lodów [30 min]***

**Celem tej części jest wprowadzenie uczestników do programu. Instruktorzy powinni się przedstawić.**

***[Slajd 1]***

Dzień dobry! Nazywam się \_\_\_\_\_\_\_\_\_ i będę waszym moderatorem podczas tej sesji.

Powiedz uczestnikom, czego mogą oczekiwać od programu:

„Będziecie się tutaj uczyć samodzielnie, a także od siebie nawzajem”.

„Moim zadaniem jest tylko zaprezentować wam różne pojęcia i koncepcje. Będziecie uczyć się nowych rzeczy SAMODZIELNIE lub od waszych przyjaciół, a po drodze dowiecie się czegoś nowego o sobie i o innych”.

„Zachęcam was, byście odpowiedzi na różne pytania szukali w Internecie lub pytali siebie nawzajem. W ten sposób uczycie się tego, jak się uczyć!”

Oczekuję od was 3 rzeczy:

* **Chęci do nauki** - Umysł nastawiony na rozwój,
* **Chęci, by dać z siebie wszystko** - Zaangażowanie,
* **Chęci do dzielenia się z innymi** - Otwartość.

DLACZEGO UCZESTNICY BIORĄ UDZIAŁ W PROGRAMIE?

* Aby nauczyć się analizować problemy i projektować możliwe rozwiązania.
* Aby nauczyć się stosować technologię AI z korzyścią dla siebie i społeczeństwa.

***[Slajd 2]***

Potrzebna jest kartka

* + Uczestnicy piszą na niej swoje imię i 3 rzeczy, które lubią. ***[5 minut]***
  + Następnie robią obchód po sali i poznają nowe osoby. Poznane osoby dopisują na kartce swoje imię i oddają głos na jedną z 3 rzeczy, którą oni najbardziej lubią. Uczestnicy mogą sobie pomagać w wypełnianiu kartek.
  + Może to potrwać od 5 do 10 minut. Każdy z uczestników zapozna się w tym czasie z 5–10 nowymi osobami. **[10 minut]**
  + To jest czas, kiedy uczestnicy mogą lepiej się poznać i odkryć, z kim mają wspólne zainteresowania.
  + Pod koniec zapytaj dowolnego uczestnika o 3 rzeczy, które lubi najbardziej, i o to, która z nich dostała najwięcej głosów. Następnie poproś, by uczestnik wskazał kogoś, kto zapisał się na jego/jej kartce. Możesz przejść wtedy do wskazanego uczestnika i zadać te same pytania. Możesz dodać warunek, by następna wskazywana osoba siedziała przynajmniej 3 stoliki dalej, aby zaangażować uczestników z różnych części sali.
  + Możesz też zapytać, kto ma najwięcej znajomych na liście. Na przykład: kto ma więcej niż 10 znajomych na swojej liście? Kto ma więcej niż 15? ... Aż nikt już nie podniesie ręki.

***[Slajd 3]***

Potencjalne pytania dodatkowe:

„Kto mi przypomni, czego się nauczyliśmy na poprzednich warsztatach?”

„Kto mi poda przykład z prawdziwego życia, dotyczący domeny AI, o której się uczymy?”

„Jak myślicie, czym się będziemy dziś zajmować?”

Pozwól 3–5 uczestnikom na udzielenie odpowiedzi, nie przerywaj dyskusji, jeśli okaże się owocna.

Chodzi o przypomnienie zdobytej wiedzy i jej zastosowania, a także o szybką rozgrzewkę i integrację uczestników na nowo po przerwie.

***[Slajd 4]***

To są rezultaty edukacyjne tych warsztatów.

[Zob. slajdy]

Dużo tu będzie samokształcenia i zadań praktycznych.

Chęć do dzielenia się, nauki i uczestnictwa jest kluczowa.

***[Slajd 5] Powtórka: Pozyskiwanie i eksploracja danych [10 min]***

Kto pamięta proces AI?

[Poczekaj na odpowiedzi uczestników.]

Dziś użyjemy języka Python do pobrania i eksploracji zbioru danych! Później przetworzymy dane, aby sprawdzić, czy nie niczego nie brakuje i czy są tam wartości odstające.

Uwaga do moderatora: Tokenizacja oznacza podzielenie dokumentu na pojedyncze słowa, czyli tokeny.

Normalizacja obejmuje takie czynności jak zmianę wszystkich słów na pisane z małej litery czy usunięcie znaków przestankowych. Lematyzacja i stemming również mogą być przeprowadzone.

### ***[Slajdy 6–18] Samokształcenie [210 min]***

**Celem tej części jest zapoznanie uczestników z importem i przetwarzaniem danych poprzez samodzielne uczenie się i praktykę.**

***[Slajd 6] Pozyskiwanie danych [60 min]***

Teraz będziecie uczyć się samodzielnie.

Wykonacie kilka ćwiczeń w notatnikach Jupyter na laptopach.

Staniecie przed różnymi wyzwaniami i zadaniami.

Być może po drodze będziecie mieli pytania. Spróbujcie odpowiedzi znaleźć w Internecie.

Zapytajcie kolegów i koleżanki. Pomagajcie sobie wzajemnie.

Jeśli mimo to nie uda wam się znaleźć odpowiedzi, zapiszcie pytania i później do nich wrócimy.

Zaczynamy!

***[Slajd 7]***

Kończymy pracę w notatnikach.

Zróbmy sobie podsumowanie!

Czego się nauczyliście? Zróbmy listę!

[Zapisz listę na tablicy]

***[Slajd 8]***

Nauczyliśmy się:

[Zob. slajdy]

[Dopasuj do tego, co powiedzieli uczestnicy]

[Przejdź przez zadania/wyzwania z uczestnikami]

Sprawdźmy, jak sobie poradziliście z zadaniami i wyzwaniami!

[Przejdź przez zadania/wyzwania i zapytaj uczestników o różne podejścia]

Uwagi do zadań/wyzwań:

1. W przypadku większości zadań istnieje więcej niż 1 możliwe podejście.

2. Jeśli niektórzy uczestnicy wykonają zadanie bardzo szybko, daj im trudniejsze wyzwanie.

3. Jeśli niektórzy uczestnicy nie potrafią wykonać wszystkich zadań, powiedz im, żeby się nie przejmowali i żeby próbowali dalej. W trakcie fazy projektu będzie czas na dzielenie się doświadczeniami i dalszy rozwój.

4. Nie dziw się, jeśli uczestnicy zaproponują metody/techniki, o których nic nie wiesz, a które oni znaleźli w Internecie.

***[Slajd 9] Kodowanie sekwencyjne [30 min]***

Gratulacje, właśnie ukończyliście wasz pierwszy notatnik. Czas teraz na grę. Jest to gra zespołowa, do której potrzebujemy grup 4-osobowych. Podzielcie się na grupy i usiądźcie razem. Za chwilę przekażę wam instrukcje.

[Przygotuj plan sali według przewodnika po zadaniach grupowych. Gdy grupy się już utworzą, wyjaśnij reguły i instrukcje gry z użyciem przewodnika po zadaniach grupowych.]

***[Slajd 10]***

Oto instrukcje:

[Zob. slajdy]

***[Slajd 11]***

Oto instrukcje:

[Zob. slajdy]

***[Slajd 12]***

Gratulacje dla wszystkich zespołów!

***[Slajd 13] Podstawy przetwarzania i wizualizacji danych [60 min]***

Przejdziemy teraz do drugiego notatnika. Drugi notatnik obejmuje podstawy statystyki, wykresy pudełkowe, histogramy i wykresy punktowe. Jak pamiętacie z poprzednich warsztatów dotyczących pozyskiwania danych, wykresy pudełkowe służą do wizualizacji wartości procentowych i rozstępu ćwiartkowego, histogramy pokazują rozkład w oparciu o rozmiary koszy, a wykresy punktowe identyfikują punkty danych oddalone od siebie. Możecie teraz zacząć pracę w drugim notatniku, zatytułowanym Podstawy przetwarzania i wizualizacji danych.

Przejdź się po sali, aby sprawdzić, czy uczestnicy mają trudności z zadaniami. Możesz objaśnić zadania, które sprawiają problemy większości uczestników. Możesz również przejść do kolejnego slajdu, jeśli większość uczestników przeszła do części bonusowej lub ma trudności z tą częścią.]

***[Slajd 14]***

Kończymy pracę w notatnikach.

Zróbmy sobie podsumowanie!

Czego się nauczyliście? Zróbmy listę!

[Zapisz listę na tablicy]

***[Slajd 15]***

Nauczyliśmy się:

[Zob. slajdy]

[Dopasuj do tego, co powiedzieli uczestnicy]

[Przejdź przez zadania/wyzwania z uczestnikami]

Sprawdźmy, jak sobie poradziliście z zadaniami i wyzwaniami!

[Przejdź przez zadania/wyzwania i zapytaj uczestników o różne podejścia]

Uwagi do zadań/wyzwań:

1. W przypadku większości zadań istnieje więcej niż 1 możliwe podejście.

2. Jeśli niektórzy uczestnicy wykonają zadanie bardzo szybko, daj im trudniejsze wyzwanie.

3. Jeśli niektórzy uczestnicy nie potrafią wykonać wszystkich zadań, powiedz im, żeby się nie przejmowali i żeby próbowali dalej. W trakcie fazy projektu będzie czas na dzielenie się doświadczeniami i dalszy rozwój.

4. Nie dziw się, jeśli uczestnicy zaproponują metody/techniki, o których nic nie wiesz, a które oni znaleźli w Internecie.

***[Slajd 16] Postępowanie z danymi błędnymi i brakującymi [60 min]***

Teraz możemy przejść do trzeciego notatnika. Ten notatnik pomoże wam zrozumieć różnicę między danymi błędnymi a odstającymi. Obejmuje on również sposób postępowania z danymi brakującymi i błędnymi. Zanim zaczniemy pracować w trzecim notatniku, przypomnijmy sobie podstawowe pojęcia. Dane błędne to dane niepoprawne, które nie powinny się znaleźć w zbiorze danych. Z kolei dane odstające to punkty danych położone daleko od innych punktów. Mogą to być wartości bardzo małe lub bardzo duże. Dane odstające wcale nie muszą być błędne. Na koniec możemy również usunąć błędy i dane brakujące ze zbioru danych, zanim go użyjemy. Ten notatnik pokaże wam, jak to zrobić. Możecie teraz uruchomić notatnik zatytułowany Postępowanie z danymi błędnymi i brakującymi.

Przejdź się po sali, aby sprawdzić, czy uczestnicy mają trudności z zadaniami. Możesz objaśnić zadania, które sprawiają problemy większości uczestników. Możesz również przejść do kolejnego slajdu, jeśli uczestnicy uzupełnili część 1.]

***[Slajd 17]***

Kończymy pracę w notatnikach.

Zróbmy sobie podsumowanie!

Czego się nauczyliście? Zróbmy listę!

[Zapisz listę na tablicy]

***[Slajd 18]***

Nauczyliśmy się:

[Zob. slajdy]

[Dopasuj do tego, co powiedzieli uczestnicy]

[Przejdź przez zadania/wyzwania z uczestnikami]

Sprawdźmy, jak sobie poradziliście z zadaniami i wyzwaniami!

[Przejdź przez zadania/wyzwania i zapytaj uczestników o różne podejścia]

Uwagi do zadań/wyzwań:

1. W przypadku większości zadań istnieje więcej niż 1 możliwe podejście.

2. Jeśli niektórzy uczestnicy wykonają zadanie bardzo szybko, daj im trudniejsze wyzwanie.

3. Jeśli niektórzy uczestnicy nie potrafią wykonać wszystkich zadań, powiedz im, żeby się nie przejmowali i żeby próbowali dalej. W trakcie fazy projektu będzie czas na dzielenie się doświadczeniami i dalszy rozwój.

4. Nie dziw się, jeśli uczestnicy zaproponują metody/techniki, o których nic nie wiesz, a które oni znaleźli w Internecie.

### ***[Slajdy 19–27] Projekt [100 min]***

**Celem tej części jest zastosowanie umiejętność zdobytych podczas sesji.**

***[Slajd 19] Projekt (część 1 z 2) [40 min]***

Brawo dla wszystkich! Dowiedzieliśmy się dzisiaj wielu fascynujących rzeczy. Sprawdźmy teraz nowe umiejętności.

Podzielcie się na 4-osobowe zespoły. Każdy zespół powinien mieć przynajmniej jeden laptop.

Gdy wszyscy będą gotowi, powiem wam, na czym będzie polegał projekt.

***[Slajd 20]***

Oto projekt!

[Zob. slajdy]

***[Slajd 21]***

Projekt podzielony jest na różne poziomy.

[Zob. slajdy]

Waszym zadaniem jest dotarcie do najwyższego poziomu!

***[Slajd 22]***

Zastanówcie się nad tymi pytaniami w trakcie planowania.

[Zob. slajdy]

Zróbcie plan, zanim zabierzecie się za projekt!

Niech każdy zespół przygotuje strategię. Podzielcie się zadaniami i nie próbujcie od razu zaczynać projektu.

***[Slajd 23] Kontrola w połowie drogi [20 min]***

Minęła połowa czasu!

Podzielcie się tym, co udało wam się do tej pory zrobić!

***[Slajd 24]***

* [Poproś uczestników, by podzielili się swoimi postępami ustnie lub pisemnie (np. używając fiszek)]
* [Zob. pytania na slajdach]
* [Udziel szybkich wskazówek i zachęć zespoły do dalszej pracy nad zadaniem!]
* [Jeśli wszystkie zespoły posuwają się do przodu zbyt wolno, można przydzielić poszczególnym zespołom różne poziomy. Nawet jeśli nie zdążą zapisać wszystkiego w kodzie, powinni potrafić opisać planowane rozwiązanie i planowany podział pracy.]

***[Slajd 25] Projekt (część 2 z 2) [40 min]***

Kontynuujcie pracę nad projektem!

***[Slajd 26]***

Przypominam: to jest zadanie obowiązkowe.

[Zob. slajdy]

***[Slajd 27]***

A to są poszczególne poziomy.

[Zob. slajdy]

***[Slajdy 28–32] Prezentacja projektu [40 min]***

***[Slajd 28]***

„Czas minął!

Teraz zaprezentujecie swoje projekty!

***[Slajd 29]***

Oto, w jaki sposób się zorganizujemy.

***[Slajd 30]***

Oto, w jaki sposób zaprezentujecie projekty:

[Zob. slajdy]

[Zwróć uwagę na cenne i ważne punkty, którymi warto podzielić się z wszystkimi.]

[Zwróć uwagę na proste sztuczki, które ułatwiają rozwiązanie pozornie skomplikowanych problemów.]

[Możesz też poprosić zespoły, by pokazały swoje dane przed przetworzeniem i po przetworzeniu, i zapytać inne zespoły, by zgadły, co zrobiono ze zbiorami danych]

***[Slajd 31]***

Gratuluję wszystkim zespołom!

***[Slajd 32]***

Omówmy teraz poszczególne projekty.

[Zob. slajdy]

***[Slajdy 33–37] Podsumowanie [30 min]***

***[Slajd 33]***

Zróbmy sobie małe podsumowanie. Powiedzcie, czego się nauczyliście!

***[Slajd 34]***

Co osobiście wydaje wam się szczególnie użyteczne?

***[Slajd 35]***

Wskażcie jedną technikę, której dziś użyliście!

***[Slajd 36]***

Jak chcecie wykorzystać umiejętności, które dziś zdobyliście?

[Zastanów się, czy istnieją zastosowanie mogące pomóc w rozwiązaniu problemów globalnych lub w bezpośrednim otoczeniu uczestników.]

[Zachęć uczestników, by uwzględnili aspekty społeczne/cele zrównoważonego rozwoju.]

***[Slajd 37]***

Sprawdźmy, czy udało nam się osiągnąć wszystkie rezultaty edukacyjne!

***[Slajd 38] Quiz [30 min]***

Czas na quiz!

Łącze do quizu: <https://create.kahoot.it/share/quiz-module-16-experience-data-data-import-and-processing/155856c1-df3c-41be-9729-9f9fd4351caa>

Jeśli natraficie na trudność, sprawdźcie, czy znajdziecie odpowiedź w Internecie.

Powszechny dostęp do Internetu sprawia, że umiejętność wyszukiwania potrzebnych informacji jest bardzo ważna.

Możecie szukać w Internecie odpowiedzi na większość pytań.

Musicie też wiedzieć, jak zweryfikować, czy znalezione odpowiedzi są prawidłowe.

Na przykład, możecie to łatwo przetestować w notatnikach Jupyter.

***[Slajdy 39–40] Refleksje [30 min]***

***[Slajd 39]***

Wszystkim należą się brawa za quiz!

Teraz podsumujmy nasze warsztaty.

Utrwalimy sobie zdobytą wiedzę i sposoby jej praktycznego zastosowania.

Notatki moderatora:

1. Poproś uczestników, by zapisali swoje refleksje. W tym czasie przejdź się po sali, upewniając się, że wykonują zadanie, i odpowiedz na wszelkie pytania lub wątpliwości.

2. Następnie pozwól im na dyskusję w zespołach i zapisanie wybranych refleksji na kolorowych kartkach.

3. Poproś 2 zespoły o podzielenie się refleksjami. Wybierz dwa dowolne zespoły, jeśli nie ma ochotników.

4. Podsumuj ogólne cele zajęć i osiągnięcia w zmaganiu z wyzwaniami.

***[Slajd 40]***

Oto pytania do refleksji.

[Zob. slajdy]

**4. Wskazówki dotyczące rozwiązywania problemów - WAŻNE**

### ***Typowe błędy/problemy sprzętowe***

|  |  |  |  |
| --- | --- | --- | --- |
|  | **BŁĘDY/PROBLEMY** | **MOŻLIWE PRZYCZYNY** | **ROZWIĄZANIE** |
| 1 | Notatnik Jupyter informuje o brakujących modułach | Moduły nie są prawidłowo zainstalowane | Sprawdź, czy moduły zostały zainstalowane za pomocą polecenia „! pip freeze” w notatniku. Zainstaluj brakujące moduły w środowisku wirtualnym. |
| 2 | Awarie środowiska wirtualnego | Konflikt zainstalowanych modułów | Odtwórz środowisko wirtualne i ponownie zainstaluj moduły |
| 3 | Awarie notatnika Jupyter | Problemy z kernelem notatnika | Zrestartuj kernel notatnika, zamknij i ponownie otwórz notatnik |

### **5. Rubryka oceny**

|  |  |  |  |
| --- | --- | --- | --- |
| Umiejętności | **Słabe** | **Średnie** | **Dobre** |
| Umiejętność importowania danych z użyciem Python | Nie potrafią importować danych bez przykładowego kodu. | Potrafią importować dane przy minimalnych podpowiedziach i z kodem przykładowym. | Potrafią importować dane bez podpowiedzi i bez przykładowego kodu. |
| Umiejętność zautomatyzowania pobierania danych | Nie potrafią zautomatyzować pobierania danych bez przykładowego kodu. | Potrafią zautomatyzować pobieranie danych przy minimalnych podpowiedziach i z przykładowym kodem. | Potrafią zautomatyzować pobieranie danych bez podpowiedzi i bez przykładowego kodu. |
| Umiejętność korzystania z percentyli | Nie potrafią korzystać z percentyli. | Potrafią korzystać z percentyli przy minimalnych podpowiedziach. | Potrafią korzystać z percentyli bez podpowiedzi. |
| Znajomość wykresów pudełkowych i histogramów | Nie rozumieją wykresów pudełkowych i histogramów | Rozumieją wykresy pudełkowe i histogramy przy minimalnych podpowiedziach. | Rozumieją i wykorzystują wykresy pudełkowe i histogramy. |
| Znajomość różnicy między błędami a wartościami odstającymi | Nie rozumieją różnicy między błędami a wartościami odstającymi | Rozumieją różnicę między błędami a wartościami odstającymi przy minimalnych podpowiedziach. | Rozumieją różnicę między błędami a wartościami odstającymi bez podpowiedzi. |
| Umiejętność postępowania z danymi błędnymi i brakującymi | Nie potrafią postępować z danymi błędnymi i brakującymi. | Potrafią postępować z danymi błędnymi i brakującymi przy minimalnych podpowiedziach. | Potrafią postępować z danymi błędnymi i brakującymi bez podpowiedzi. |

### **6. Polecane lektury**

### **Wprowadzenie do języka Python na potrzeby nauki o danych (podstawowe wprowadzenie do języka Python i tablic NumPy)** <https://www.datacamp.com/courses/intro-to-python-for-data-science>

### **Praca z plikami w Python**

<https://realpython.com/working-with-files-in-python/>

### **Użytkowanie programu wget** <http://cmdlinetips.com/2011/08/how-to-download-a-file-or-directory-using-wget/>

### <https://www.programcreek.com/python/example/83386/wget.download>

### <https://pypi.org/project/wget/>

### **Użytkowanie biblioteki Pandas**

<https://pandas.pydata.org/pandas-docs/stable/10min.html>

<https://www.tutorialspoint.com/python_pandas/index.htm>

### **Użytkowanie biblioteki Matplotlib**

<https://realpython.com/python-matplotlib-guide/>

### **Wprowadzenie do eksploracji danych w Python**

<https://realpython.com/working-with-files-in-python/>

### **Oczyszczanie danych za pomocą Python i Pandas: wykrywanie wartości brakujących**

<https://towardsdatascience.com/data-cleaning-with-python-and-pandas-detecting-missing-values-3e9c6ebcf78b>

### **Sposoby wykrywania i usuwania wartości odstających**

<https://towardsdatascience.com/ways-to-detect-and-remove-the-outliers-404d16608dba>

### **Wprowadzenie do eksploracji danych w Python**

<https://realpython.com/working-with-files-in-python/>