

# Kurs Jamovi – Zadania

Jacek Matulewski

WWW: <https://jacekmatulewski.fizyka.umk.pl/dydaktyka/jamovi/index.html>

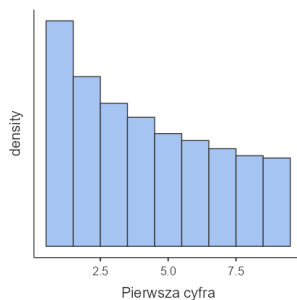
## Zadanie 1. BMI

1. Ze strony <https://cedarcreek.umn.edu/research/experiments/e267/dataset/data-age-gender-height-weight-activity-levels-each-household> pobierz dane dotyczące wagi, wzrostu, wieku i aktywności fizycznej osób (kopia dostępna pod adresem <https://jacekmatulewski.fizyka.umk.pl/dydaktyka/jamovi/dane/bmi.txt>).
2. Dla uproszczenia zmień nazwę pliku na *bmi.txt*.
3. Wczytaj te dane do Jamovi.
4. Uporządkuj i typy i przelicz wartości niektórych zmiennych:
  - a. nazwę zmiennej *Member\_Age\_Orig* zmień na *Wiek* i ustaw typ danych jako całkowity (wiek jest podawany w pełnych latach);
  - b. nazwę zmiennej *Member\_Gender\_Orig* zmień na *Płeć* i zmień nazwy poziomów (można je edytować w panelu *Zmienna*): 1 na Mężczyzna, 2 na Kobieta;
  - c. dodaj zmienną wyliczaną *Wzrost\_cm*, która zawiera wartości ze zmiennej *HEIGHT* przeliczone ze stóp na centymetry (formuła  $=HEIGHT * 30.48$ );
  - d. dodaj zmienną wyliczaną *Waga\_kg*, która zawiera wartości ze zmiennej *WEIGHT* przeliczone z funtów na kilogramy (formuła  $=WEIGHT * 0.45359237$ );
  - e. nazwę zmiennej *Mod\_act* zmień na *Liczba godzin umiarkowanego wysiłku* i ustaw typ danych jako całkowity zostawiając skalę ciągłą;
  - f. analogicznie nazwę zmiennej *Vig\_act* zmień na *Liczba godzin energicznego wysiłku* i ustaw typ danych jako całkowity.
5. Dodaj zmienną wyliczaną o nazwie *BMI* zawierającą wartości *Body Mass Index* (waga w kilogramach podzielona przez wzrost w metrach). Do obliczenia kwadratu służy w Jamovi formuła  $"^2"$ . **Uwaga!** W tej chwili dysponujemy wzrostem w centymetrach.
6. Zrób histogram wartości BMI.
7. Czy średnia i mediana BMI mieszczą się w zakresie wagi prawidłowej (18,5 - 25)? Jeżeli nie, sprawdź, czy BMI jest istotnie różne od wartości 21,75.
8. Przygotuj dwa filtry uwzględniający jedynie osoby z nadwagą ( $BMI \geq 25$  i  $BMI < 30$ ) i otyłe ( $BMI \geq 30$ )?
9. Sprawdź, czy istnieje zależność między BMI a wiekiem oraz umiarkowaną i energiczną aktywnością fizyczną (korelacje).
10. Przygotuj wykres rozszutu pokazujący zależność BMI od wieku (można podzielić na grupy ze względu na płeć).
11. Sprawdź, czy średnie BMI różnią się istotnie u kobiet i mężczyzn (pamiętaj o sprawdzeniu normalności rozkładu, aby wybrać właściwy test).
12. Analogicznie porównaj BMI osób z niską i wysoką aktywnością fizyczną (podziel na grupy według liczby godzin energicznej aktywności fizycznej: grupa o niskiej aktywności niech zawiera osoby o liczbie godzin mniejszej od średniej o jedno odchylenie standardowe, a wysokiej aktywności – większe o jedno odchylenie).
13. Wyeksportuj wyniki do pliku PDF.

Rozwiązanie dostępne jest w pliku *bmi.omv*.

## Zadanie 2. Prawo Benforda

1. Uruchom Jamovi i usuń zmienne A, B i C.
2. W menu *Dane, Wstaw*, wybierz pozycję *Na końcu*. Pojawi się pytanie o liczbę wierszy. Wstaw 1000. Zwróć uwagę, że po użyciu *Zmienne, Wstaw, Na końcu* to pytanie się nie pojawia. Ta czynność pozornie nie przyniesie efektu (nie powstanie nowa zmienna – to chyba błąd), ale tabela zostanie rozszerzona do tysiąca wierszy. I to jest dla nas najważniejsze.
3. Następnie dodaj zmienną wyliczaną o nazwie N zawierającą kolejny numer wiersza (można do tego użyć formuły  $=\text{ROW}()$ ).
4. Następnie dodaj zmienną wyliczaną o nazwie Wartości zawierającą kolejne potęgi dowolnej liczby  $q$ . Dla  $q = 2$  będą to 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096, 8192, 16384, itd. (formuła  $=N^2$ )
5. Alternatywnie do 1-4 można wygenerować plik CSV zawierający serię tysiąca kolejnych wyrazów ciągu geometrycznego korzystając z ChatuGPT i wczytując go do Jamovi.
6. Naszym celem jest utworzenie dwóch zmiennych wyliczanych zawierających pierwszą i ostatnią cyfrę wartości elementów ciągu geometrycznego. Niestety Jamovi nie pozwala funkcji, które dają gotowe wartości (także funkcji dla tekstu, które zwracałyby pierwszą i ostatnią literę ciągu). Dlatego musimy poradzić sobie dzięki matematyce.
  - a. Ostatnią cyfrę liczby możemy otrzymać jako resztę z dzielenia przez 10. Operator zwracający taką resztę to  $\%$  (modulo).
  - b. Pierwszą cyfrę otrzymamy dzieląc przez 10 do potęgi równej liczbie cyfr liczby pomniejszonej o jeden i zaokrąglając w dół do liczby całkowitej. Dla 2025 będzie to  $2025/10^{4-1} = 2025/100 = 2.015$ . Po zaokrągleniu w dół (funkcja `FLOOR`) otrzymamy 2.
  - c. Potrzebną w poprzednim punkcie liczbę cyfr można uzyskać korzystając z logarytmu o podstawie 10 (funkcja `LOG10`) po zaokrągleniu wyniku w górę (ponownie funkcja `FLOOR` – dzięki temu będzie od razu pomniejszona o 1).
7. Sprawdź rozkład pierwszych i ostatnich cyfr dla  $q = 2$  (wykreśl histogramy).



Dlaczego częstość występowania pierwszych cyfr nie jest jednorodny (maleje wraz ze wzrostem wartości cyfry)? Sprawdź tę zależność dla  $q = 3$ , dla ciągu Fibonacciego lub dla liczebności miast w Polsce

(źródło: [https://pl.wikipedia.org/wiki/Dane\\_statystyczne\\_o\\_miastach\\_w\\_Polsce](https://pl.wikipedia.org/wiki/Dane_statystyczne_o_miastach_w_Polsce)

lub Baza Danych Lokalnych GUS <https://bdl.stat.gov.pl/BDL/start>).

Odpowiedź znajdziesz w [https://youtu.be/EoNQ\\_mC5WQ8](https://youtu.be/EoNQ_mC5WQ8).

Rozwiązanie dostępne jest w pliku *benford.omv*.

### Zadanie 3. Rzuty dwoma kostkami do gry

1. Usuń zmienne A, B i C, a następnie korzystając ze „sztuczki” opisanej w poprzednim zadaniu dodaj do tabeli 10 wierszy.
2. Dodaj dwie zmienne wyliczane o nazwach `Kostka1` i `Kostka1`, w których losowane będą liczby oczek wyrzucanych oboma kostkami (formuła `=FLOOR (UNIF (1, 7) )`).
3. Sprawdź rozkład wartości obu zmiennych na histogramie.
4. Następnie zwiększ liczbę wierszy do 100, a potem do 1000, obserwując histogramy.
5. Dla tysiąca elementów rozkłady obu zmiennych powinny być już wystarczająco bliskie do rozkładu jednorodnego.
6. Dodaj zmienną wyliczaną o nazwie `Suma` będącą sumą losowanych wartości w zmiennych `Kostka1` i `Kostka2`.
7. Sprawdź rozkład jej wartości. Czy jest jednorodny?
8. Korzystając z testu Shapiro-Wilka sprawdź, czy ten rozkład jest zgodny z rozkładem normalnym.
9. Powtórz te czynności dla nowej zmiennej wyliczanej będącej iloczynem liczby oczek z obu kostek.
10. Następnie zastąp rozkład jednorodny rozkładem normalnym (Gausa), co oznacza zastąpienie funkcji `UNIF` funkcją `NORM`. Zwróć jednak uwagę na inne znaczenie obu argumentów w nowej funkcji. Zamiast minimalnej i maksymalnej wartości są to teraz średnia i odchylenie standardowe. Sprawdź normalność rozkładu zmienionych zmiennych `Kostka1` i `Kostka2` oraz ich sumy i iloczynu.

Rozwiązanie dostępne jest w pliku *kostki.omv*.