Wskazówki dotyczące interpretacji parametrów z próby:

1. Średnią arytmetyczną i odchylenie standardowe (inne parametry też) liczy się po to, aby rozkłady empiryczne zastąpić wyidealizowanymi rozkładami teoretycznymi (modelami), które w PRZYBLIŻENIU odzwierciedlają swoje empiryczne źródło, nie tylko ułatwiając, ale wręcz umożliwiając dokonywanie operacji i ocen, które bez tego są praktycznie niewykonalne.
2. Średnia arytmetyczna jest liczbą, dla której suma różnic między nią a każdym pomiarem wynosi zero;
   a. Średnia arytmetyczna jest liczbą, dla której suma kwadratów różnic między nią a każdym pomiarem jest najmniejsza (oczywiście suma kwadratów różnic musi być większa od zera);
   b. Główną wadą jest to, że wszystkie wartości są tak samo ważne i mają identyczny wpływ na wynik.
   c. Duży wpływ wartości skrajnych na wartość średniej.
   d. Identyfikujemy wartości odstające, by ocenić ich wpływ na średnią i asymetrię rozkładu.
   e. Oceniamy przyczynę istnienia wartości odstających - Błąd pomiaru? Przypadek? Rzeczywiste zjawisko?
3. Mediana (wartość środkowa) jest liczbą, dla której suma bezwzględnych różnic między nią a każdym pomiarem jest minimalna
4. Dominanta (moda, wartość najczęstsza) to wartość, która występuje najczęściej. W rozkładach dyskretnych liczona jest wprost z definicji. W rozkładach ciągłych przyjmuje się, że dominanta to wartość, w której rozkład przyjmuje wartość maksymalną.
   a. Jeśli rozkład ma klika maksimów lokalnych, to mówimy, że rozkład jest wielomodalny.
   b. Wielomodalność wskazuje na konieczność analizy w celu wyodrębnienia grup obiektów rozróżnialnych niewidoczną wcześniej cechą.
5. W rozkładach symetrycznych średnia, mediana i dominanta są równe.
   a. Asymetrię oceniamy biorąc różnicę między średnią i dominantą. Jeżeli ta różnica jest ujemna to mówimy o asymetrii lewostronnej wskazującej na istnienie skrajnych wyników po lewej stronie (małe wartości) próby. Jeżeli ta różnica jest dodatnia to mówimy o asymetrii prawostronnej wskazującej na istnienie skrajnych wyników po prawej stronie (duże wartości) próby.
   b. Współczynnik asymetrii $A_s > 0$- asymetria prawostronna (prawostronna skośność):
      i. dominanta < mediana < średnia
      ii. prawa strona (prawy ogon rozkładu) jest dłuższy
      iii. im większe A_s tym mocniejsza prawostronna skośność
   c. Współczynnik asymetrii $A_s = 0$ – rozkład symetryczny,
      i. średnia = mediana = dominanta
      ii. obie strony rozkładu są podobne (lustrzane odbicie)
   d. Współczynnik asymetrii $A_s < 0$ – asymetria lewostronna (lewostronna skośność):
      i. dominanta > mediana > średnia
      ii. lewa strona (lewy ogon rozkładu) jest dłuższy
      iii. im mniejsze A_s tym mocniejsza lewostronna skośność

6. Wariancja służy do wyznaczenia odchylenia standardowego. Jest to średnia kwadratów odchyleń pomiarów od średniej z tych pomiarów. Odchylenie standardowe jest miarą rozproszenia próby.
   a. Odchylenie standardowe wykorzystywane jest do wyznaczenia wartości odstających.
   b. Odchylenie standardowe wykorzystywane jest do oceny zgodności rozkładu empirycznego z rozkładem normalnym (Reguła 3-sigma) – wartości typowych1 jest ok. 68% a wartości typowych2 ok. 95,5%
   c. Wartości typowe 1 są to wartości, które mieszczą się w przedziale wartości typowych postaci:
      $$\langle średnia - odchylenie\ standardowe; średnia + odchylenie\ standardowe \rangle$$
   d. Wartości typowe 2 są to wartości, które mieszczą się w przedziale wartości typowych:
      $$\langle średnia - 2 * odchylenie\ standardowe; średnia + 2 * odchylenie\ standardowe \rangle$$
   e. Procent wartości typowych w próbie wyznaczamy dzieląc liczbę wartości w przedziale wartości typowych przez liczebność próby. Procent ten porównujemy z wartościami dla rozkładu normalnego (punkt a.)
   f. Wartości odstające są to wartości, które nie leżą w przedziale wartości typowych typu 2.
      i. Średnia jest wrażliwa na wartości odstające tzn. wartości odstające wpływają na wartość średniej co może prowadzić do nieprawidłowych wniosków
      ii. Mediana i dominanta są odporne na wartości odstające
7. Odchylenie przeciętne, podobnie jak odchylenie standardowe jest miarą rozproszenia próby. Obecnie prawie nie jest wykorzystywane.
8. Kwantyle czyli (Kwartyle, Centyle, Decyle itp.) stosuje się do oceny zakresu typowych wartości w populacji.
   a. Kwartyl pierwszy dzieli zbiorowość na dwie części w ten sposób, że 25% jednostek zbiorowości ma wartości cechy niższe bądź równe kwartylowi pierwszemu, a 75% równe bądź wyższe od tego kwartyla
   b. Kwartyl drugi = Mediana – 50% jednostek zbiorowości ma wartości cechy niższe bądź równe kwartylowi drugiemu, i 50% równe bądź wyższe od tego kwartyla
   c. Kwartyl trzeci – analogicznie, tylko podział jest 75% do 25%

d. Centyle powstają jak kwartyle, ale dokonujemy podziału co 1%, czyli w skali centylowej 50 oznacza Medianę. Przy ocenie rozwoju niemowląt stosuje się skalę centylową np. jeśli dziecko ma wagę poniżej 10 centyli, to wskazuje na problemy zdrowotne powodujące niedożywienie dziecka.
e. Decyle powstają przez podział próby co 10%
a. W oparciu o kwartyle można wyznaczyć również miarę rozproszenia czyli odchylenie ćwiartkowe.
b. W oparciu o kwartyle i odchylenie ćwiartkowe można wyznaczyć wartości odstające, by potem ocenić przyczynę ich istnienia w próbie.

9. Współczynnik zmienności to miara zróżnicowania wartości.
   a. Interpretacja wartości współczynnika zmienności:
      i. poniżej 25% – bardzo mała zmienność;
      ii. przedział 25-25% – przeciętna zmienność;
      iii. przedział 45-100% – silna zmienność;
      iv. powyżej 100% – bardzo intensywna zmienność.
   b. Na podstawie współczynnika zmienności: można analizować statystyczne zmienne w obszarze różnych populacji, czy cech.

Liczebność próby	n	
Pomiary, dane, wartości próby	$x_i, \quad i = 1, \dots, n$	
Rozstęp	$x_{max} - x_{min}$	x_{max}-wartość największa w próbie x_{min}-wartość najmniejsza w próbie
Dominanta (moda)	D	Wartość, która występuje najczęściej
Średnia	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	
Wariancja populacji	$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$	Stosujemy, jeśli wnioski dotyczą tylko zebranych danych (populacji)
Odchylenie standardowe populacji	$s = \sqrt{s^2}$	Stosujemy, jeśli wnioski dotyczą tylko zebranych danych (populacji)
Wariancja z próby	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$	Stosujemy, jeśli wnioski z danych stosujemy do szerszej grupy (populacja)
Odchylenie standardowe z próby	$s = \sqrt{s^2}$	Stosujemy, jeśli wnioski z danych stosujemy do szerszej grupy (populacja)
Typowe 1	$\langle \bar{x} - s, \bar{x} + s \rangle$	procent typowych 1 w rozkładzie normalnym wynosi 68%
Typowe 2	$\langle \bar{x} - 2s, \bar{x} + 2s \rangle$	procent typowych 2 w rozkładzie normalnym wynosi 95,5%
Wartości odstające	Wartości z próby nienależące do zbioru typowe 2	
Klasyczny współczynnik zmienności	$v = \frac{s}{\bar{x}}$	
Klasyczny współczynnik asymetrii	$A_s = \frac{\bar{x} - D}{s}$	$A_s = 0 \ gdy \ s = 0$
Parametry pozycyjne		
Mediana = kwartyl drugi	M	wartość środkowa, w próbie uporządkowanej od najmniejszej do największej wartości
Kwartyl pierwszy	Q_1	leży w 1/4 próby

Kwartyl trzeci	Q_3	leży w 3/4 próby
Rozstęp międzykwartylowy	$Q_3 - Q_1$	
Odchylenie ćwiartkowe	$Q = \frac{Q_3 - Q_1}{2}$	Pełni rolę analogiczną do odchylenia standardowego
Kwartylowe typowe 1	$\langle M - Q, M + Q \rangle$	
Kwartylowe typowe 2	$\langle Q_3 - 3Q, Q_3 + 3Q \rangle$	
Kwartylowe wartości odstające	Wartości z próby nienależące do zbioru kwartylowe typowe 2	
Kwartylowy współczynnik zmienności	$v = \frac{Q}{\overline{M}}$	
Kwartylowy współczynnik asymetrii	$As = \frac{(Q_3 - M) - (M - Q_1)}{2(Q_3 - Q_1)}$	$A_s = 0 \; gdy \; Q_3 = Q_1$
Mieszany współczynnik asymetrii	$A_s = 3\frac{\bar{x} - M}{s}$	$A_s = 0 \; gdy \; s = 0$