

ANALIZA DANYCH ILOŚCIOWYCH W PIGUŁCE

dr n. społ. Anna Dzielska

ZANIM PRZEJDZIESZ DO ANALIZ
DOBRE PRAKTYKI

USTAL CO CIĘ INTERESUJE? JAKIE ZJAWISKO CHCESZ ZBADAĆ?

Zapoznaj się z literaturą dotyczącą interesującego cię zjawiska:

- Co piszą o tym inni badacze?
- Co do tej pory zostało zbadane?
- Jakie są wyniki tych badań?
- Co mówi o tym zjawisku teoria lub teorie naukowe?
- Na jakie pytania mogę jeszcze odpowiedzieć?

SFORMUŁUJ PROBLEM BADAWCZY I PYTANIA BADAWCZE

Możesz sformułować jeden lub kilka problemów badawczych, a w ich obrębie po kilka szczegółowych pytań badawczych.

Dobre pytanie badawcze powinno być:

- precyzyjne
- krótkie i niezbyt złożone (lepiej abyś sformułował więcej pytań niż w jednym zawierał kilka problemów do rozwiązania)
- możliwe do zbadania biorąc po uwagę dostępną wiedzę teoretyczną i narzędzia, którymi dysponujesz

Pytania badawcze najczęściej dotyczą poszukiwania różnic między grupami lub badania jaki jest związek między zmiennymi

POSTAW HIPOTEZY BADAWCZE

Hipoteza badawcza – założenie, które wymaga sprawdzenia, weryfikacji.

Próba odpowiedzi na pytanie, które wynika z problemu badawczego.

Może wskazywać kierunek zależności (hipoteza kierunkowa) lub nie przewidywać kierunku zależności (hipoteza niekierunkowa).

Zdecyduj, który rodzaj hipotez (kierunkowa czy niekierunkowa) pasuje do twojego badania.

POSTAW HIPOTEZY BADAWCZE C.D.

Hipoteza zerowa – mówi o braku zależności

Na przykład:

- Kierunkowa: obraz ciała chłopców nie jest lepszy niż obraz ciała dziewcząt
- Niekierunkowa: obraz własnego ciała nie jest związany z płcią

Hipoteza alternatywna – mówi o istnieniu zależności

Na przykład:

- Kierunkowa: obraz ciała chłopców jest lepszy niż obraz ciała dziewcząt
- Niekierunkowa: obraz własnego ciała różni się ze względu na płeć

USTAL W JAKI SPOSÓB CHCESZ BADAĆ INTERESUJĄCE CIĘ ZJAWISKO

- kogo będziesz badał (osoby badane – próba),
- w jaki sposób (metody badań),
- jakiego narzędzia lub narzędzi użyjesz aby zgromadzić materiał badawczy (kwestionariusz ankiety, scenariusz wywiadu strukturyzowanego, formularz obserwacji i in.)

STWÓRZ NARZĘDZIE/ NARZĘDZIA BADAWCZE

- Przygotowując narzędzie badawcze kieruj się postawionymi problemami i pytaniami badawczymi.
- Korzystaj z gotowych testów lub przygotuj wszystkie pytania samodzielnie.
- Dobierz pytania tak abyś mógł wyczerpująco przeanalizować i opisać to co chcesz zbadać.

ZBIERZ DANE

- Pamiętaj o tym, że powinieneś uzyskać zgodę na udział w badaniu od każdej osoby, która przystępuje do badań.
- Jeśli badasz osoby niepełnoletnie, najpierw poproś o zgodę ich rodziców, następnie zapytaj badanych czy zgadzają się na udział. Jeśli badasz dzieci poniżej 13 roku życia zgoda nie jest konieczna ale kierując się etyką badań warto o nią zapytać 😊. Dobrowolność, nacisku i presji = rzetelne odpowiedzi.

PRZYGOTUJ DANE DO ANALIZY

- Wprowadź odpowiedzi do bazy danych.
- Zanim zaczniemy wprowadzać dane do programu SPSS powinieneś zakodować pytania, które znajdowały się w kwestionariuszu. Każde pytanie ponumeruj lub nazwij (nazwy zmiennych nie mogą się powtarzać), a odpowiedziom przypisz kody (np. przez ponumerowanie odpowiedzi kolejnymi cyframi).

PRZYKŁAD:

Pytanie: Gdzie mieszkasz?

Nazwa zmiennej: miejsce_zamieszkania

Kategorie odpowiedzi	Kod (numer, który wprowadzisz do bazy danych)
Na wsi	1
W małym mieście	2
W dużym mieście	3

PRZEJDŹ DO ANALIZY ZEBRANEGO MATERIAŁU BADAWCZEGO

UWAGA:

- W rozdziale metodologicznym konieczne zamieść akapit, w którym opisziesz jakich metod analizy danych używałeś. Najlepiej wypisz te metody odnosząc się do kolejnych pytań badawczych/ hipotez, które sformułowałeś w swojej pracy.
- Opisz na jakich założeniach się opierałeś aby potwierdzić lub obalić postawione hipotezy badawcze. W tym miejscu podaj np. informację o tym, że wartość „p”, która pozwala uznać, że wyniki są istotne statystycznie wynosi $p < 0,05$.
- Napisz również z jakiego oprogramowania do analizy danych korzystałeś.

ANALIZA DANYCH

Odpowiednio sformułowane pytania badawcze oraz hipotezy pozwolą ci uporządkować proces analizy.

Dobierz do nich metody, za pomocą których będziesz analizował dane.

PODSTAWOWE STATYSTYKI OPISOWE

Na początek przeprowadź proste analizy, które pozwolą opisać twoją próbę i jej cechy oraz zmienne analizowane w badaniu.

- **Miary występowania, tzw. „częstości”** – rozkład odpowiedzi na pytania udzielone przez badanych: procent (procent „ważnych” z raportu SPSS) i n-liczba badanych, która odpowiada temu procentowi)
- **Miary tendencji centralnej** – np. średnia (suma wszystkich wartości zmiennej podzielona na liczbę przypadków – średnia arytmetyczna), mediana (wartość środkowa rozkładu), modalna (in. dominanta – wartość zmiennej, która występuje najczęściej w zbiorze danych)
- **Miary zmienności** – wariancja (miara rozproszenia wyników wokół średniej - wysoki wynik mówi o tym, że wyniki są bardzo zróżnicowane i w małym stopniu skupione wokół średniej), lub odchylenie standardowe (mówi o tym jakie jest przeciętne zróżnicowanie wartości zmiennej od średniej tej zmiennej)

ANALIZA ZALEŻNOŚCI – WYBRANE METODY

KORELACJE

Wyniki analizy korelacji r-Pearsona pomagają uzyskać odpowiedź na pytanie czy pomiędzy dwoma zmiennymi istnieje związek liniowy. Warunkiem jest aby obie zmienne miały postać ciągłą a nie kategoryczną.

Siła tego związku mówi o tym jaki jest stopień współzmienności – nie mówi o zależności przyczynowo - skutkowej. Oprócz poziomu istotności statystycznej ważna jest uzyskana w analizie wartość współczynnika korelacji r-Pearsona.

KORELACJE C.D.

Interpretacja wartości współczynnika r-Pearsona (wg. Bedyńska, Cypryńska, 2012)

- 0-0,3 – brak korelacji lub bardzo słaba korelacja
- 0,3-0,5 – umiarkowana korelacja
- 0,5-0,7 – silna korelacja
- 0,7-1 – bardzo silna korelacja

Wartość współczynnika (ujemna lub dodatnia) również na kierunek związku między zmiennymi:

- korelacja ujemna (znak „-” przed wartością współczynnika) – wysokim wartościom jednej zmiennej towarzyszą wysokie wartości drugiej zmiennej lub niskim wartościom jednej zmiennej towarzyszą niskie wartości drugiej zmiennej
- korelacja dodatnia (znak „+” przed wartością współczynnika) - niskim wartościom jednej zmiennej towarzyszą wysokie wartości drugiej zmiennej

TABELE KRZYŻOWE (TEST CHI-KWADRAT)

Za pomocą tabel krzyżowych czyli zestawienia wartości dwóch zmiennych można przeprowadzić test chi-kwadrat, dzięki czemu uzyskamy odpowiedź na pytanie czy istnieje zależność między tymi zmiennymi.

Wynik istotny statystycznie ($p < 0,05$) informuje o istnieniu zależności.

PORÓWNANIE ŚREDNICH

Test t Studenta dla prób zależnych:

Stosowany do porównania średnich wartości między dwiema grupami (według zmiennej, która posiada 2 cechy, np. jak różni się średnia wartość w skali poczucia własnej wartości między chłopcami a dziewczętami)

Jednoczynnikowa analiza wariancji ANOVA:

Jeśli zmienna posiada więcej niż dwie cechy to do porównania średnich używamy jednoczynnikowej analizy wariancji ANOVA (np. chcemy zbadać jak różni się średnia ocena obrazu własnego ciała w grupie studentów uniwersytetów, politechnik i wyższych szkół artystycznych)

ANALIZA REGRESJI LINIOWEJ

Dzięki niej możemy przewidzieć jak zmienią się wartości zmiennej zależnej pod wpływem zmiany wartości zmiennych niezależnych (predyktorów). Za pomocą regresji liniowej budujemy model linowy tej zależności.

Możemy badać wpływ jednej zmiennej niezależnej – wtedy mówimy o regresji jednozmiennowej lub więcej niż jednej zmiennej niezależnej – wtedy jest to regresja wielozmiennowa.

TU WARTO ZAJRZEĆ

KSIĄŻKI:

Książki:

Analiza danych krok po kroku – od wprowadzania danych, przez proste do złożonych analiz statystycznych z wykorzystaniem programu IBM SPSS, na przykładach z nauk społecznych.

W kolejnych rozdziałach znajdują się odpowiedzi na niemal wszystkie pytania, które nurtują zarówno początkujących jak i bardziej zaawansowanych analityków.

- **Bedyńska S, Cypryańska M, Statystyczny drogowskaz 1. Praktyczne wprowadzenie do wnioskowania statystycznego, Sedno, SWPS, 2012**
- **Bedyńska S, Cypryańska M, Statystyczny drogowskaz 2. Praktyczne wprowadzenie do analizy wariancji, Sedno, SWPS, 2013**

TUTORIALE:

Na youtube można znaleźć mnóstwo filmów, których autorzy krok po kroku pokazują jak analizować dane i jak interpretować uzyskane wyniki (output z programu SPSS) 😊

- **Przykłady kursów/ tutoriali:**

Kurs SPSS dla studentów psychologii od coding.pl

- https://www.youtube.com/watch?v=jX4_zmYEGgU&list=PLg1_IjaxUikIV1QcBVHgiQMRTjSCwMhiL

Prezentacja różnych etapów pracy z SPSS:

- <https://www.youtube.com/user/wicherekstoklosy/videos>

POWODZENIA!!!

UWAGA:

Niniejsze opracowanie potraktuj jako wprowadzenie do swoich poszukiwań.

Zanim przejdziesz do analiz zapoznaj się ze szczegółowymi zasadami ich prowadzenia (min. warunkami jakie muszą spełnić dane) – skorzystaj ze wskazanej literatury ☺

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2$$

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{\sum_{i=1}^N x_i}{N}$$