

Radosław Poniak  
Białystok

**Propozycja zasad prezentowania wyników analiz statystycznych  
w publikacjach naukowych  
z zakresu dziejów społeczno-gospodarczych**

Charakterystyczna dla polskiej historiografii mała popularność analiz statystycznych jest zjawiskiem dobrze znanym, na które wielokrotnie zwracano już uwagę<sup>1</sup>. W porównaniu z historykami z Zachodu krajowi badacze znacznie rzadziej sięgają po metody kwantytatywne, a jeśli już to czynią, to bardzo często w sposób najprostszy z możliwych, ograniczając się do podania procentów lub średnich arytmetycznych. Zaawansowane techniki statystyczne napotkać można w polskich pracach tylko sporadycznie, a ich ewentualne zastosowanie nie pociąga zwykle za sobą naśladownictwa w publikacjach innych historyków. W konsekwencji więc pojedyncze przypadki skutecznego wykorzystania metod statystycznych nie wyznaczają nowych wzorców, a znaczna część środowiska ignoruje takie publikacje właśnie z powodu użycia w nich niezrozumiałych dla ogółu metod kwantytatywnych.

Nieczęste sięganie po techniki statystyczne oraz wyspowy charakter ich wykorzystania sprawiają, iż w naszej praktyce badawczej nie wykształcił się powszechnie uznawany i stosowany zestaw reguł określających zasady opracowywania baz danych, prowadzenia analiz oraz przedstawiania zestawień statystycznych w publikacjach naukowych. Nawet w środowiskach charakteryzujących się stosunkowo wysokim stopniem integracji i dysponujących podręcznikami

---

<sup>1</sup> Piotr Guzowski, Radosław Poniak, *Miejsce badań kwantytatywnych we współczesnej historiografii polskiej*, „Roczniki Dziejów Społecznych i Gospodarczych” 73, 2013, s. 243–255.

wyznaczającymi standardy postępowania badawczego<sup>2</sup> brakuje jednoznacznych i powszechnie podzielanych reguł prezentowania wyników. Konsekwencją takiej sytuacji są częste przypadki analiz prowadzonych według założeń niezgodnych z regułami uznawanymi przez większość środowiska, co prowadzi do braku porównywalnych rezultatów, wyników budzących duże wątpliwości, uzyskanych za pomocą niejasnych procedur i raportowanych w sposób niejednoznaczny albo wręcz niezrozumiały. Pojawia się też niebezpieczeństwo, że wraz z większą elastycznością w dopuszczanym na danym polu stosowaniu metod badawczych, opracowywaniu danych i formach przedstawiania wyników rośnie prawdopodobieństwo wprowadzenia do obiegu naukowego rezultatów o wątpliwej jakości, niespójnych, a czasem wręcz będących owocem celowych manipulacji<sup>3</sup>. Prowadzi to do marnotrawienia wysiłku poszczególnych badaczy, niemożności stosowania technik określanych zbiorczo mianem metaanalizy<sup>4</sup>, braku integracji środowiska wokół wspólnych tematów i wyników.

Wobec opisanej powyżej sytuacji oczywista staje się więc potrzeba ujednoczenia procedur i wypracowania podstawowych standardów prezentacji danych statystycznych. Niniejszy artykuł powstał w celu zainicjowania dyskusji nad tym zagadnieniem oraz wskazania kilku rozwiązań wartych jak najszybszego upowszechnienia. Inspirację do jego napisania stanowiły dyskusje prowadzone z wieloma badaczami stosującymi metody kwantytatywne. W jakimś stopniu stanowi on próbę oddania nie tylko moich poglądów, ale też opinii wyrażanych przez co najmniej część środowiska polskich demografów historycznych i historyków społeczno-gospodarczych. W niektórych przypadkach wskazać zresztą można na osoby, które część z przedstawionych tu myśli prezentowały na konferencjach lub seminariach czy podczas polemik dotyczących publikacji ukazujących się w polskiej historiografii. W związku z tym swoją rolę postrzegam tutaj przede wszystkim jako sekretarza, który próbuje w miarę systematyczny sposób oddać toczącą się od pewnego czasu debatę, a nie głóściciela nowych, dotąd nieznanych poglądów. Mam nadzieję, że lista zaprezentowanych tu problemów i propozycji,

---

<sup>2</sup> Za przykład mogą tu posłużyć demografowie historyczni zgrupowani wokół regularnie się zbierającego Zespołu Demografii Historycznej PAN i korzystający z dwóch podręczników: Ireny Gieysztorowej, *Wstęp do demografii staropolskiej*, Warszawa 1976 oraz Cezarego Kukli, *Demografia Rzeczypospolitej przedrozbiorowej*, Warszawa 2009.

<sup>3</sup> John P.A Ioannidis, *Why Most Published Research Findings Are False*, „PLoS Medicine” 2005, No. 8: e124, DOI:10.1371/journal.pmed.0020124.

<sup>4</sup> Dobrym przykładem zastosowania metaanalizy na polu demografii historycznej może być szeroko dyskutowany artykuł dotyczący zasadności uznawania europejskiego modelu małżeństwa za jeden z czynników warunkujących rozwój gospodarczy: Tracy Dennison, Sheilagh Ogilvie, *Does the European marriage pattern explain economic growth?*, „Journal of Economic History” 74, 2014, issue 3, s. 651–693. W języku polskim na znaczenie tego tekstu zwracał uwagę Radosław Poniat, *Nowe spojrzenie na europejski model małżeństwa*, „Historyka. Studia Metodologiczne” 45, 2015, s. 193–199.

w żadnym wypadku niewyczerpująca zagadnienia, będzie stanowiła nie tylko inspirację do dalszych dyskusji, ale też zachętę do poprawy standardu tekstów wykorzystujących metody kwantytatywne, w tym również artykułów publikowanych na łamach PDP.

Choć podane poniżej propozycje wynikają z obserwacji zjawisk zachodzących w polskiej historiografii i każda z nich poparta może być konkretnymi przykładami niewłaściwego zastosowania i prezentowania metod statystycznych, które odnaleźć można w pracach krajowych badaczy, czytelnik nie odnajdzie tu dokładnych namiarów bibliograficznych na publikacje zawierające opisywane błędy. Sądzę, że wskazywanie na poszczególnych autorów byłoby nieuczciwe w sytuacji, gdy podobne pomyłki popełnia też wiele innych osób, a samo występowanie niewłaściwych praktyk nie wynika ze złej woli danego badacza, ale z niewystarczająco wysokich standardów obowiązujących w ramach dyscypliny. Muszę zresztą przyznać, że sam nie w każdej publikacji stosowałem się do podawanych tu sugestii. Czasem wynikało to z braku wiedzy, czasem lenistwa, a często z nacisku recenzentów i redaktorów. Dlatego też czytelnikowi będzie musiało jedynie wystarczyć moje stwierdzenie, że omawiane tu błędy występują w pracach polskich historyków. Zapewne zresztą bez trudu on sam będzie mógł wskazać na choć niektóre ich przykłady.

### **Podawanie liczebności badanej zbiorowości**

Informacja na temat liczby przypadków, na których podstawie dokonano analiz statystycznych, jest konieczna dla oceny znaczenia wyprowadzanych wniosków oraz porównywania rezultatów. Nawet mało ze statystyką obeznany badacz intuicyjnie pojmuje przecież różnicę między formułowaniem sądów na podstawie tylko kilku albo tysięcy przypadków. W związku z tym unikanie przez historyków informacji o rzeczywistej wielkości podstawy swych obliczeń musi być oceniane podobnie jak wybiórcze przywoływanie źródeł lub ukrywanie nazwy zespołu archiwalnego, z którego one pochodzą.

Niestety, w pracach historyków często odnaleźć można przypadki, gdy dane na temat liczby obserwacji są podawane nieprecyzyjnie lub w ogóle nie występują. Choć sytuacje skrajne, polegające na całkowitej rezygnacji z informowania czytelnika o wielkości badanej zbiorowości<sup>5</sup>, spotykane są dość sporadycznie, wciąż wskazać można na przypadki, gdy mimo pozornego podania takiej informacji rzeczywista liczebność populacji pozostaje nieznana. Zaskakująco często

---

<sup>5</sup> Pojęcie to dotyczy nie tylko kwestii takich jak liczba ludności badanej parafii, ale też zdarzeń demograficznych, liczby wzmianek w źródłach dotyczących wybranego zjawiska, wielkości podstawy źródłowej.

autorzy tekstów naukowych poprzestają na jednorazowej wzmiance o liczbie badanych obiektów, zaś w dalszych analizach pojawiają się już wyłącznie zestawienia procentowe lub średnie arytmetyczne. Gdy wyliczenia takie dotyczą całej zbiorowości, czytelnik może podejrzewać, że oparte są one na wszystkich obiektach, ale często okazuje się to nieprawdą. Część elementów zostaje bowiem usunięta z analizy z powodu braku danych, niejasności zapisu lub niemożności zakwalifikowania do konkretnej kategorii. Czasem pominięcia takie dotyczyć mogą znacznego odsetka wszystkich obserwacji, ale czytelnik albo nie jest o tym w ogóle informowany, albo musi zadowolić się wzmianką w stylu: „po usunięciu części przypadków”.

Do jeszcze większych przekłamań prowadzić może sytuacja, gdy badana zbiorowość jest dzielona na grupy albo dane zjawisko analizowane jest w podziale na podokresy lub wybrane kategorie. Procedury te nieuchronnie prowadzić muszą do znacznego spadku liczby obserwacji, na których podstawie wyciągane są bardziej szczegółowe wnioski. W skrajnych przypadkach takie postępowanie skutkować może podawaniem wskaźników sezonowości przy starannym ukrywaniu faktu, że liczone są na podstawie tylko kilkunastu zdarzeń. W celu uniknięcia takiego niebezpieczeństwa konieczne jest oczekiwanie od autorów precyzyjnego podawania liczby obserwacji w każdej wyróżnionej grupie lub podokresie.

### Szersze wykorzystywanie miar statystycznych

Jeśli autorzy polskich prac historycznych decydują się na sięgnięcie po miary statystyczne opisujące rozkład badanej populacji, niemal zawsze sprowadza się to jedynie do wykorzystania średniej arytmetycznej. Nie negując zalet tej miary, pamiętać jednak należy o jej ograniczeniach.

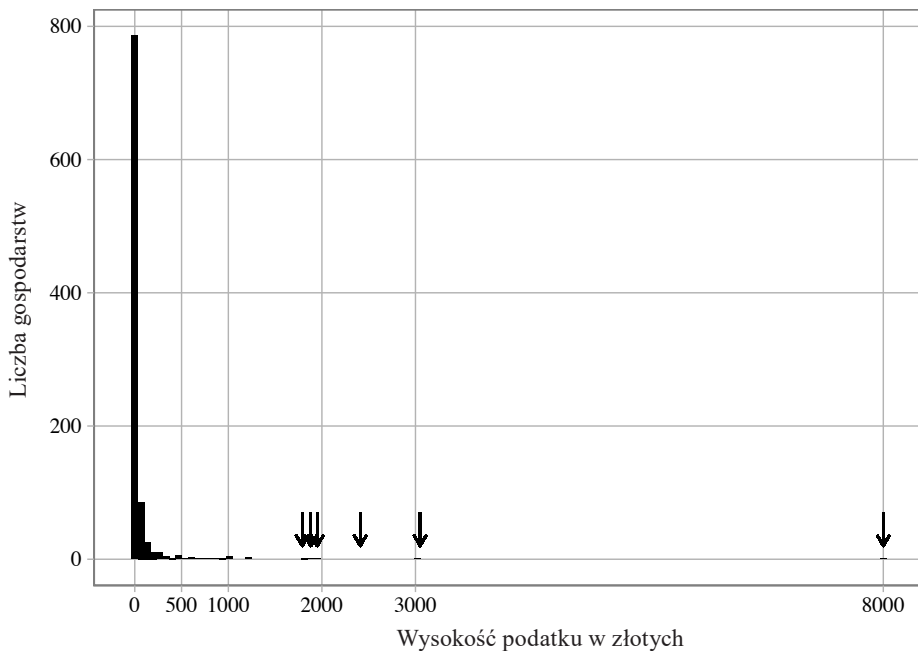
Po pierwsze, wskazać można na sytuacje, gdy średnia okazuje się miarą nieadekwatną, dającą wyniki ukrywające rzeczywisty rozkład danych. Im bardziej analizowana zbiorowość odbiega od rozkładu normalnego, przestaje być symetryczna i pojawiają się w niej obserwacje znacznie odstające wielkością od pozostałych, tym mniej dokładnie będzie ona opisywana przez średnią arytmetyczną. Dobrego przykładu takiego zjawiska dostarczyć mogą dane zaprezentowane na wykresie 1. Przedstawiają one wysokość podatku płaconego w roku 1794 przez gospodarstwa domowe grodzieńskich chrześcijan<sup>6</sup>. Nawet pobieżna obserwacja

---

<sup>6</sup> Był to podatek przeznaczony na spłatę kontrybucji nałożonej na populację miasta przez wojsko rosyjskie. Ponieważ każda głowa gospodarstwa musiała zapłacić kwotę stanowiącą 1% wartości swego majątku, podatek ten pozwala na dość precyzyjny opis struktury majątkowej mieszkańców Grodna. Zagadnienie to zostało szerzej omówione w artykule: Radosław Poniat, *Zróżnicowanie majątkowe mieszkańców Grodna w 1794 roku*, „Klio. Czasopismo Poświęcone Dziejom Polski i Powszechnym” 2015, nr 1 (32), s. 83–109.

danych pozwala z łatwością dostrzec, że przeważają tu gospodarstwa obciążone niewielkimi zobowiązaniami. Na 954 wszystkich gospodarstwach domowych tylko 148, czyli około 15%, wpłacić musiało kwoty przekraczające 50 złp. Jednak średnia arytmetyczna, wynosząca 60 złp, wskazałaby na wyższą przeciętną wysokość podatku! Rezultat taki jest konsekwencją występowania w badanej populacji gospodarstw płaćcych sumy znacznie odbiegające od wartości przeciętnych. Najbogatsze spośród nich, na wykresie niemal niewidoczne i dlatego zaznaczone za pomocą strzałek, stanowiły wprawdzie nieznaczny procent ogółu, ale ze względu na wysokość wpłacanych podatków zdecydowanie zawyżyły otrzymany rezultat. W takiej sytuacji, w przypadku analiz struktur własności bardzo częściej, wykorzystanie średniej arytmetycznej prowadzi do wyników nieoddających rzeczywistości i wręcz wprowadzających badaczy oraz czytelników w błąd. Znacznie użyteczniejszą miarą okazuje się tu chociażby mediana, mniej podatna na zakłócenie przez wartości odstające. W odniesieniu do opisywanej tu populacji wynosi ona zaledwie 6 złp i zdecydowanie lepiej oddaje strukturę majątkową grodzieńskich chrześcijan.

Wykres 1. Wysokość podatków płaconych przez chrześcijańskich mieszkańców Grodna w 1794 roku (w złp)



Źródło: *Spis mieszkańców Grodna z 1794 r.*, opr. Jerzy Urwanowicz, [w:] *Grodno w XVIII wieku. Miasto i ludność*, red. Andrzej Woltanowski, Jerzy Urwanowicz, Białystok 1997, s. 61–171.

Wydaje się, że podany tu przykład dobrze ilustruje zagrożenia związane z opieraniem się wyłącznie na średniej arytmetycznej i wskazuje na korzyści płynące z sięgania po inne miary statystyczne. Autorzy powinni więc jak najczęściej obok średniej podawać też medianę lub inne miary tendencji centralnej. Postępowanie takie należy zalecać szczególnie w przypadku, gdy wartości poszczególnych miar statystycznych znacznie różnią się między sobą.

Należy też pamiętać, że sama średnia lub mediana informują nas tylko o jednym parametrze rozkładu – punkcie centralnym. Z perspektywy badacza równie ważne są miary statystyczne wskazujące na jego zróżnicowanie. Niestety, w krajowej literaturze informacje dotyczące odchylenia standardowego albo miar odpornych takich jak rozstęp ćwiartkowy czy MAD pojawiają się niezwykle rzadko. Stan taki nie tylko uniemożliwia wyliczanie przedziałów ufności (co omówiono poniżej), ale też prowadzić może do błędnej interpretacji otrzymanych wyników.

Za przykład takiego zjawiska posłużyć może analiza przeciętnego zaludnienia miast w dwóch galicyjskich cyrkulach w 1824 roku<sup>7</sup>. Średnie arytmetyczne wskazują, że miasta w cyrkule stanisławowskim i złoczowskim charakteryzowały się podobnym zaludnieniem, wynoszącym przeciętnie 2624 mieszkańców w pierwszym z nich i 2627 w drugim. Wniosek taki byłby jednak błędny. Choć średnie są w obydwu przypadkach niemal identyczne, to powiązane z nimi odchylenia standardowe różnią się bardzo wyraźnie. Miara ta wynosi 1639 w cyrkule stanisławowskim i 3472 w złoczowskim<sup>8</sup>. Oznacza to, że druga jednostka administracyjna charakteryzowała się znacznie większym zróżnicowaniem zaludnienia miast, występowały w niej obok siebie ośrodki bardzo małe i całkiem duże, podczas gdy w cyrkule stanisławowskim dysproporcje między ośrodkami były znacznie mniejsze. Obserwację taką potwierdza analiza rozkładów przedstawionych na wykresie 2. Widoczne na nim czarne punkty prezentują zaludnienie miast, zaś romby wskazują na położenie średnich arytmetycznych.

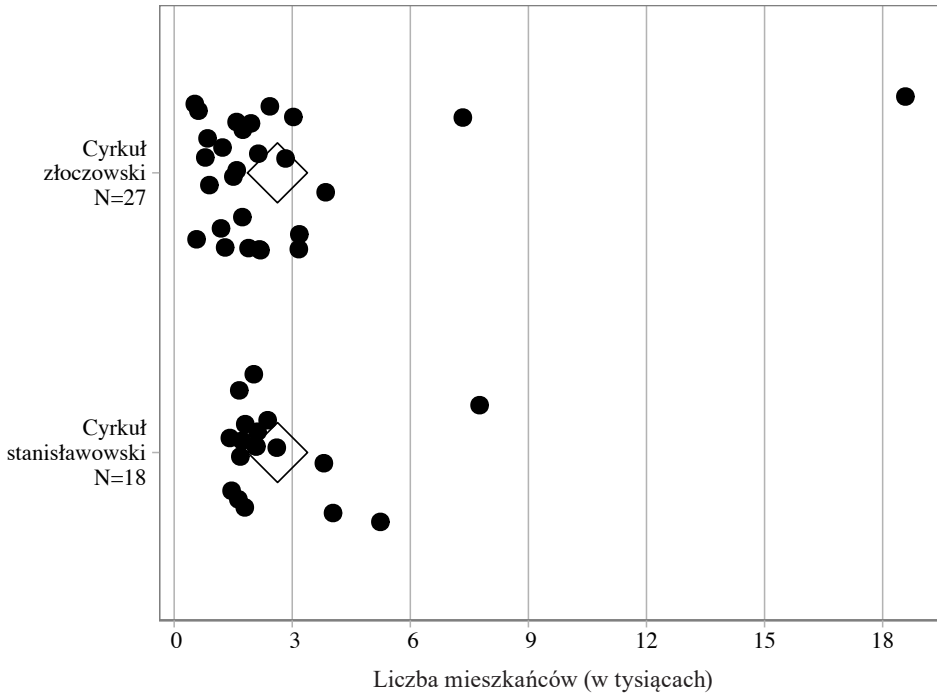
Podany przykład dobrze ilustruje niebezpieczeństwa związane z podawaniem przez badaczy jedynie średniej arytmetycznej i pomijaniem innych miar statystycznych, w tym przede wszystkim tych opisujących rozkład analizowanej zbiorowości. Poprzestawanie na zaprezentowaniu tylko jednej miary tendencji centralnej, nawet jeśli w adekwatny sposób opisuje ona badaną populację, nieuchronnie prowadzić będzie do niemożności porównywania rezultatów, nadmiernej redukcji przekazu lub wręcz wprowadzania czytelników w błąd. Dlatego też średniej arytmetycznej lub medianie zawsze powinny towarzyszyć nie

---

<sup>7</sup> Łukasz Jewuła, *Galicyjskie miasta i miasteczka oraz ich mieszkańcy w latach 1772–1848*, Kraków 2013. Dane pochodzą z Aneksu dołączonego do książki.

<sup>8</sup> Co ważne, w podanym przykładzie również analiza median nie wskazałaby na istnienie znacznych różnic w strukturze zaludnienia miast. W obydwu cyrkulach mediany są do siebie zbliżone i wynoszą 2048 mieszkańców w cyrkule stanisławowskim oraz 1745 w złoczowskim.

Wykres 2. Zaludnienie miast cyrkułów stanisławowskiego i złoczowskiego w 1824 roku



Źródło: Łukasz Jewuła, *Galiczyjskie miasta i miasteczka oraz ich mieszkańcy w latach 1772–1848*, Kraków 2013, s. 173–245.

tylko informacja o liczebności zbiorowości, ale też miary takie jak odchylenie standardowe lub rozstęp ćwiartkowy<sup>9</sup>. Dobrze widziane byłoby także podawanie wartości maksymalnych i minimalnych, kwartyli oraz (w przypadku niektórych rozkładów) skośności.

### Ujednolicenie zapisu miar statystycznych

Interpretację i dalsze wykorzystywanie wyników analiz statystycznych ułatwiłoby także przyjęcie powszechnego standardu zapisu otrzymywanych wartości. Od autorów należy oczekiwać stosowania uniwersalnych i jednoznacznych oznaczeń liczby obserwacji (N), mediany (Me), odchylenia standardowego (S), minimum (Min) i maksimum (Max), kwartyli (Q1, Q2, Q3), rozstępu (R)

<sup>9</sup> Wybór właściwej miary wynikać musi z cech badanej zbiorowości. Im bardziej jej rozkład różni się od normalnego, tym na znaczeniu zyskuje nie tylko mediana, ale też powiązane z nią rozkład ćwiartkowy lub MAD (odchylenie bezwzględne od mediany).



i rozstępu ćwiartkowego ( $R_q$ ). Zapis taki pozwalałby na szybkie i niemylące czytelnika podawanie miar statystycznych na przykład w formie: „w roku średnia arytmetyczna zaludnienia miast w cyrkule złoczowskim wynosiła 2624 mieszkańców ( $N = 27$ ,  $S = 3472$ ,  $Me = 1745$ ,  $Min = 524$ ,  $Max = 18580$ )”. Przy prezentowaniu takich miar oraz zestawień procentowych autorzy winni być także ostrzegani przed podawaniem zbyt wielu cyfr po przecinku. W sytuacji, gdy dokonywane przez historyków analizy statystyczne bazują często na materiale źródłowym o bardzo zróżnicowanej wiarygodności, podawanie ułamków z dokładnością do tysięcznych może wprowadzać czytelników w błąd. Trzeba też wymagać przestrzegania konwencji, zgodnie z którą w tabelach niewystępowanie zjawiska zaznacza się za pomocą pauzy, brak informacji kropką, zaś zero rezerwuje się dla sytuacji, gdy zjawisko wprawdzie istniało, ale o natężeniu zbyt małym, aby dało się je oddać w tabeli.

### Prezentacja błędów standardowych

Od autorów tekstów naukowych należy oczekiwać informowania czytelników o wielkości błędów standardowych wyliczanych w pracy miar statystycznych. Ma to znaczenie przy określaniu istotnych różnic między średnimi lub wskazywaniu, które z zaobserwowanych prawidłowości są w niewielkim stopniu obciążone ryzykiem przypadkowego ich uzyskania z populacji o odmiennej charakterystyce. Równie ważna jest tu możliwość wykorzystywania wiedzy o błędach standardowych w dalszych analizach, w tym przede wszystkim badaniach porównawczych oraz metaanalizie. Niestety, w polskiej historiografii dane tego typu pojawiają się bardzo sporadycznie. Stan taki musi ulec zmianie. Nawet jeśli autorzy nie chcą lub nie potrafią sami przeprowadzać różnorodnych testów istotności lub określać przedziałów ufności, wciąż powinni dawać odbiorcom szansę na samodzielne wyliczenie takich statystyk.

W przypadku średniej arytmetycznej lub proporcji najprostszym sposobem informowania czytelników o błędach standardowych może być podawanie liczby obserwacji w analizowanych grupach oraz (w przypadku średnich) odchyłeń standardowych. Już tak podstawowe miary powinny pozwolić odbiorcy zaznajomionemu z metodami kwantytatywnymi na prowadzenie samodzielnych analiz i szerokie włączenie uzyskanych rezultatów do własnych badań.

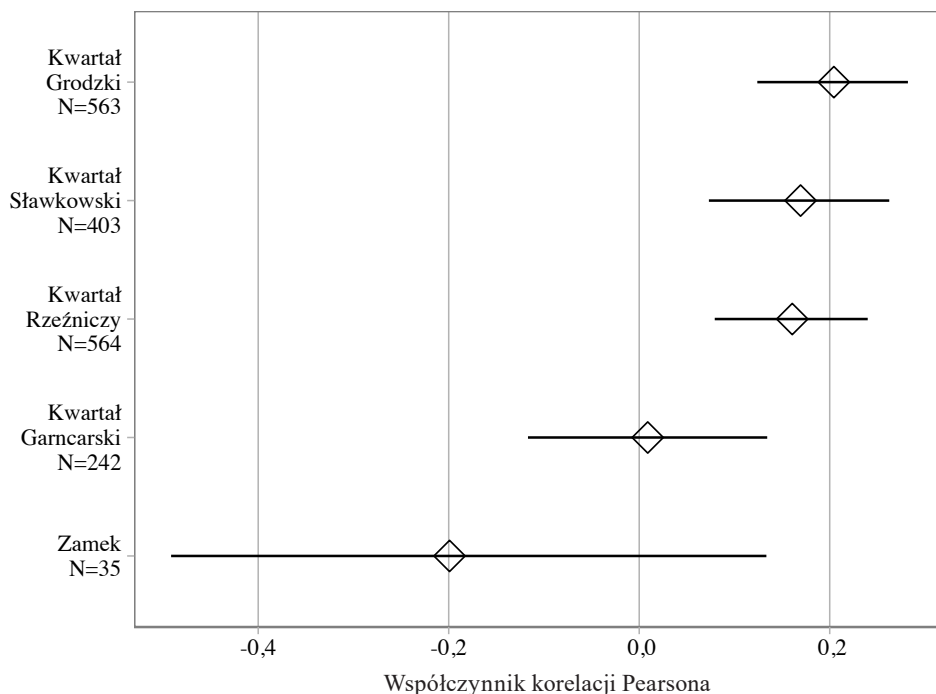
Nieco trudniejsze może być informowanie o błędach standardowych<sup>10</sup> w odniesieniu do korelacji lub regresji. W takich wypadkach konieczność

---

<sup>10</sup> Wydaje się, że informowanie o samej istotności jest znacznie gorszym rozwiązaniem niż podawanie przedziałów ufności i w żadnym przypadku nie może go w pełni zastąpić. Badania wskazują zresztą, że w celu celniejszej interpretacji wyników najlepiej byłoby istotność całkowicie zastąpić przedziałami ufności – por. Melissa Coulson, Michelle Healey, Fiona Fidler, Geoff



Wykres 3. Korelacja między liczbą dzieci a liczbą czeladzi w krakowskich gospodarstwach domowych w 1795 roku



Źródło: Archiwum Narodowe w Krakowie, IT 551.

informowania o wielkości błędu lub przedziałach ufności spadać musi na autora publikacji. Co ważne, znaczenie takich miar nie sprowadza się tylko do umożliwienia czytelnikom analizy istotności uzyskanych wyników. Równie ważne jest tu chociażby pełniejsze porównywanie znaczenia współczynników korelacji lub poszukiwanie zdarzeń odbiegających w sposób istotny od wyliczonego trendu. Przykład pierwszego z takich zastosowań zaprezentowany został na wykresie 3. Przedstawia on związek między liczbą dzieci a liczbą czeladzi w gospodarstwach domowych zarejestrowanych w tak zwanym spisie pruskim ludności Krakowa z 1795 roku<sup>11</sup>. Współczynnik korelacji Pearsona wyliczony został oddzielnie dla

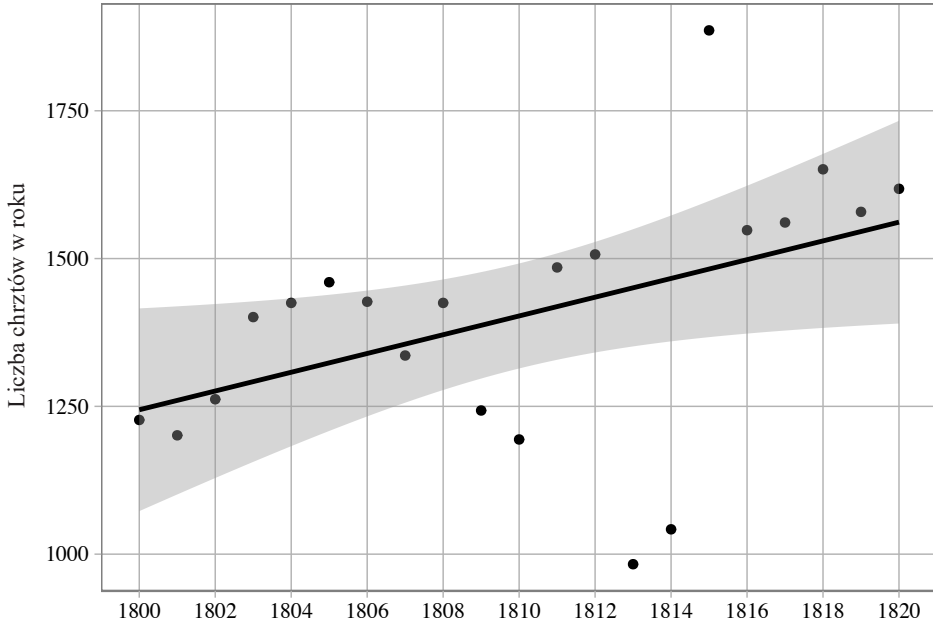
Cumming, *Confidence intervals permit, but do not guarantee, better inference than statistical significance testing*, „Frontiers Psychology” 2010, 1:26, DOI: 10.3389/fpsyg.2010.00026.

<sup>11</sup> Spis ten obejmował jedynie Kraków w obrębie murów miejskich oraz Wawel, a pomijał, administracyjnie do miasta włączone, przedmieścia i okoliczne miejscowości – por. Radosław Ponią, *Śłużba domowa w miastach na ziemiach polskich od połowy XVIII do końca XIX wieku*, Warszawa 2014, s. 119–121.

każdego z czterech kwartałów miasta oraz Zamku Królewskiego. Uzyskane wartości, oddane na wykresie za pomocą rombów, wyraźnie wskazują na istnienie słabego dodatniego powiązania między badanymi zmiennymi. Jedyńm wyjątkiem jest tu zamieszkały przez niewielką i dość specyficzną populację zamek oraz kwartał garncarski, gdzie współczynnik korelacji jest bliski zeru. Na wykresie oddane też jednak zostały za pomocą odcinków 95% przedziały ufności współczynników korelacji, które wskazują na brak ich istotności w odniesieniu do zamku oraz kwartału garncarskiego. Oczywiście, o zjawisku takim można by czytelników poinformować także za pomocą zwykłej istotności, bez korzystania z przedziałów ufności. Jednak zalety płynące z używania przedziałów nie ograniczają się jedynie do samej analizy istotności. Równie ważne jest tu wskazanie na różnice między istotnymi współczynnikami. Gdyby podanym tu korelacjom towarzyszyła tylko wiadomość o ich istotności, czytelnicy mogliby sądzić, że powiązanie między liczbą dzieci a czeladzią jest najmocniejsze w kwartale grodzkim, zaś wyraźnie słabsze na przykład w kwartale sławkowskim. Konstatacja taka mogłaby następnie posłużyć do rozważań nad różnicami w sytuacji ekonomicznej obydwu kwartałów, która skutkować mogła w biedniejszym z nich zastępowaniem kosztownego zatrudniania czeladzi przez darmową pracę własnych dzieci... Takich uwag nie potwierdza jednak obserwacja przedziałów ufności. Wynika z niej, że między badanymi kwartałami nie ma istotnej różnicy, a wyższa wartość współczynnika korelacji w przypadku grodzkiego nie uprawnia nas do stwierdzeń o jego specyfice.

Przykład korzyści, jakie płynąć mogą z podawania przedziałów ufności w wyniku analizy regresji, zaprezentowany został na wykresie 4. Widoczne na nim dane pochodzą z dobrze demografom historycznym znanej książki Jana Baszanowskiego. Do analizy wybrane zostały gdańskie chrzty z lat 1800–1820. Jednak obok samych wartości rocznych, oddanych za pomocą punktów, widoczna tu jest też linia trendu wyliczona za pomocą regresji liniowej dla szeregów czasowych. Wskazuje ona na następujący w badanym okresie wzrost liczby chrztów, choć niektóre lata dość znacznie odbiegały od tej prawidłowości. Zaznaczone na wykresie szare pole, oddające 95% przedział ufności współczynnika regresji, pozwala na wskazanie, kiedy dochodziło do takich istotnych różnic od trendu. Czytelnik łatwo dostrzeże, że przez większość badanego okresu roczna liczba chrztów nieznacznie tylko od trendu odbiegała i mieściła się w przedziale ufności. Lata specjalne wiążą się zaś przede wszystkim z kolejnymi kampaniami wojennymi i ich demograficznymi konsekwencjami. Bez zastosowania przedziałów ufności konstatacja taka wcale nie musiałaby być oczywista.

Wykres 4. Liczba chrztów w Gdańsku w latach 1800–1820



Źródło: Jan Baszanowski, *Przemiany demograficzne w Gdańsku w latach 1601–1846 w świetle tabel ruchu naturalnego*, Gdańsk 1995, s. 353–354.

### Ujednolicone grupowanie danych

Choć podany tu wymóg wydawać się może bezsporny i trudno by było wskazać na badacza, który by się z nim nie zgodził, zaskakująco często napotkać można prace, których autorzy grupują dane w tylko w sobie właściwy sposób, odbiegający znacznie od rozwiązań przyjmowanych przez innych autorów. W konsekwencji opracowania dotyczące identycznych zagadnień mogą zawierać zestawienia statystyczne całkowicie nieporównywalne. Zbyt często zdarza się, że grupowania nie są czynione z myślą o dalszej analizie czy przyszłej syntezie rezultatów, ale raczej o wygodzie danego autora. Ignorując istniejącą literaturę i dotyczące analizowanego zagadnienia teorie, liczni historycy wolą albo powtarzać określenia pochodzące ze źródła, albo, co gorsza, tworzą klasyfikacje własne, nieoparte głębszym namysłem. Szczególnie wyraźnie zjawisko takie dostrzec można w pracach dotyczących struktury zawodowej i społecznej, niemal z reguły nieodnoszących się do powszechnie uznanych międzynarodowych klasyfikacji

zawodów<sup>12</sup>, ale od problemu takiego niewolne są też publikacje poświęcone zagadnieniom demograficznym czy gospodarczym. Wystarczy tu wskazać na przypadki niezgodnych z powszechną praktyką grupowań wieku, które uniemożliwiają szersze wykorzystywanie danych zaczerpniętych z wielu prac dotyczących populacji parafialnych.

W takiej sytuacji konieczne jest ciągle przypominanie i wymaganie od autorów, aby do grupowania obserwacji wykorzystywali standardy już przyjęte w literaturze, a po nowe sięgali tylko w razie istotnej przyczyny, która zawsze musi być dokładnie wyjaśniona. Podobnie niezbędne jest podawanie precyzyjnych kryteriów klasyfikowania obserwacji do analizy oraz poszczególnych wyliczeń.

### Uwzględnianie wymiaru przestrzennego prowadzonych analiz

Trudno wyobrazić sobie wyczerpujące studium dotyczące kwestii takich jak: mobilność przestrzenna, struktury osadnicze czy funkcjonowanie rynków bez analizy ich wymiaru przestrzennego. Fakt ten był dobrze znany dawnym pokoleniom historyków, z których wielu miało zresztą geograficzne wykształcenie lub stale współpracowało z przedstawicielami tej dyscypliny<sup>13</sup>. Na Zachodzie mimo przejściowego kryzysu zainteresowanie aspektami przestrzennymi w badaniu przeszłości jest wciąż podtrzymywane, a ostatnio wręcz przeżywa okres ożywienia wpisującego się w obejmujący wiele dyscyplin tak zwany *spatial turn*<sup>14</sup>.

Niestety, krajowi badacze nazbyt często o znaczeniu geografii w swych analizach zapominają. Zjawisko to można dobrze prześledzić, przeglądając zawartość polskich czasopism historycznych w poszukiwaniu elementów kartograficznych, bez których trudno przecież wyobrazić sobie opis przestrzeni. Pomijając artykuły autorstwa archeologów, ostatnie trzy dekady charakteryzują się spadkiem liczby map zamieszczanych w czołowych periodykach, a pamiętać przecież trzeba, że i we wcześniejszych dziesięcioleciach nie były one szczególnie często spotykane<sup>15</sup>.

---

<sup>12</sup> Wskazać tu można chociażby na klasyfikacje HISCO i HISCLASS – zob. Marco H.D. van Leeuwen, Ineke Mass, Andrew Miles, *HISCO. Historical International Standard Classification of Occupations*, Lueven 2002; Marco H.D. van Leeuwen, Ineke Mass, *HISCLASS. A Historical International Social Class Scheme*, Lueven 2011. Obydwie klasyfikacje są nadal rozwijane i modyfikowane. Ich aktualne wersje odnaleźć można na stronie [www.hisma.org](http://www.hisma.org).

<sup>13</sup> Wojciech Wrzosek, *Historia – kultura – metafora. Powstanie nieklasycznej historiografii*, Wrocław 1995, s. 100–102.

<sup>14</sup> *The spatial turn. Interdisciplinary perspectives*, red. Barney Warf, Santa Arias, London 2009.

<sup>15</sup> Opinię taką potwierdza analiza bazy danych obejmującej artykuły opublikowane w latach 1951–2010 na łamach PH, PDP, RDSG oraz RH – zob. P. Guzowski, R. Poniat, *Miejsce badań kwantytatywnych* [1]. Baza ta stanowi cyfrowy aneks do artykułu, który pobrać można ze strony PDP.

Stan taki trudno zrozumieć w sytuacji, gdy na ostatnie lata przypada gwałtowny rozwój programów przeznaczonych do przetwarzania i wizualizacji danych geograficznych (GIS), co daje historykom szansę na coraz szersze uwzględnianie w swych pracach przestrzennych aspektów badanych zjawisk<sup>16</sup>.

Wydaje się, że jedynie nacisk ze strony środowiska, uosobionego przede wszystkim przez recenzentów prac naukowych i redakcje czasopism, może doprowadzić do poprawy sytuacji. Od badaczy podejmujących zagadnienia związane z przestrzenią należy oczekiwać wzięcia tego czynnika pod uwagę w prowadzonych analizach oraz korzystania z graficznych prezentacji w postaci map lub planów. Zlekceważenie takiego wymogu, na przykład w artykule dotyczącym rynku matrymonialnego lub funkcjonowania przymusu parafialnego, powinno stanowić podstawę do odrzucenia recenzowanego tekstu. Być może powinno się wprowadzić okres przejściowy, gdy dopuszczane wciąż będą wizualizacje, a w ostateczności też analizy nie w pełni satysfakcjonujące recenzentów, ale autorzy powinni być zachęceni do jak najszybszego wzbogacenia swych prac o opis przestrzennego aspektu badanych zjawisk.

### Upublicznianie baz danych

Z perspektywy długofalowych projektów badawczych oraz przyszłych syntez nie można też zapominać o potrzebie udostępniania baz danych. Choć praktyka taka pozostaje w sprzeczności z instynktami większości badaczy, którzy charakteryzują się naturalną skłonnością do ochrony wykorzystywanych źródeł i boją się utraty kontroli nad wynikami własnej pracy, to w dłuższej perspektywie byłaby ona korzystna nie tylko dla całego środowiska, ale też poszczególnych autorów. Szerszy dostęp do zbiorów danych pozwoliłby nie tylko na przyspieszenie badań i wzbogacenie ich o elementy komparatystyczne, ale też oznaczałby częstsze cytowanie osób opracowujących takie zasoby. W wielu wypadkach skutkowałby też uniknięciem powtarzania już raz przez kogoś wykonanej pracy. Ponieważ metody kwantytatywne najlepiej sprawdzają się na dużych zbiorach danych, a przy ich analizie znaczenie mają zagadnienia takie jak reprezentatywność i dobór przestrzenny obserwacji, dzielenie się danymi wiązałaby się z szansą na poszerzenie zestawu pytań badawczych podejmowanych przez polskich historyków. Trzeba zresztą pamiętać, że znaczna część ważnych osiągnięć światowej demografii historycznej i historii gospodarczej opierała się właśnie na wielkich bazach danych opracowanych przez duże zespoły badaczy. Wydaje się, że wobec

---

<sup>16</sup> Obok wyspecjalizowanych aplikacji, z których część jest darmowa, wskazać tu też trzeba na pakiety przeznaczone do wykorzystywania danych przestrzennych zaimplementowane w standardowych programach do analizy statystycznej takich jak: SPSS, STATA, R.

słabości systemu finansowania nauki w Polsce oddolna współpraca historyków może stanowić choć częściowe remedium na brak takich zespołów.

Uznanie powyższych argumentów powinno zachęcać badaczy do udostępniania jeśli nie całych baz, to choć wyczerpujących zestawień informacji z nich pochodzących. Ponieważ materiały takie ze względu na swą obszerność nie są zazwyczaj drukowane, powinny one przyjmować formę cyfrowych aneksów do artykułów lub monografii. Forma dzisiejszych czasopism naukowych, od których wymaga się istnienia stron internetowych albo wręcz przejścia w wersję elektroniczną, pozwala na bardzo łatwe wzbogacanie ich o różnego rodzaju materiały dodatkowe. Zwłaszcza w przypadku wydawania źródeł redakcje powinny wymagać dołączania do nich wersji zdigitalizowanych<sup>17</sup>. Także publikacje o bardziej analitycznym charakterze zyskałyby znacznie dzięki dodawaniu do nich cyfrowych aneksów. Konsekwencją takiej praktyki byłoby też oczywiście wymaganie od osób korzystających z takich materiałów wyraźnego powoływania się w przypisach na autorów baz i traktowania zasobów tego typu jako integralnych części danej publikacji.

Przedstawione tu propozycje, skoro wynikają z indywidualnych doświadczeń autora, nie muszą oddawać perspektywy innych badaczy pracujących na niejednakowym materiale źródłowym i funkcjonujących w odmiennych realiach instytucjonalnych. Sądzę jednak, że dyskusja i ich ewentualne przyjęcie dawałyby szansę na zwalczenie części patologii trapiących polskie badania nad dziejami społeczno-gospodarczymi i dawałyby nadzieję na szybszy rozwój dyscypliny. Bez takich zmian trudno sobie wyobrazić powstanie nowych syntez i prowadzenie szeroko zakrojonych studiów porównawczych. Ujednolicenie metod analizy i prezentacji danych powinno też zwrócić uwagę historyków na zespołowy charakter ich wysiłku, który służyć ma przecież nie tylko pisaniu indywidualnych, autorskich monografii, ale też wspólnemu odkrywaniu przeszłości, gdzie pojedyncze prace stanowią winny element wspólnych projektów badawczych.

---

<sup>17</sup> W przypadku wydawania źródeł masowych powinny to być przede wszystkim pliki łatwo czytane przez powszechnie dostępne programy bazodanowe lub arkusze kalkulacyjne.

---

**A Proposal for the Rules Concerning the Presentation of the Results  
of Statistical Analyses in Academic Publications  
within the Socio-Economic History**

**Summary**

The article presents some proposals that might contribute towards improving quality while results of statistical analyses are presented in Polish publications within history and related disciplines. According to the author nowadays there are no clear standards determining the rules of such presentations, which often leads to a limited comparability of results and makes it impossible to apply the techniques of data meta-analysis, and in some cases also leads to serious methodological errors. The proposals included in the text consists of the following postulates:

- 1) to inform the reader of the number of the populations and subpopulations investigated;
- 2) to make a greater use of statistical measures that describe the distribution of the analysed communities;
- 3) to standardise the convention of recording such measures;
- 4) to inform the reader about the size of the standard errors;
- 5) to apply the commonly accepted rules concerning data grouping;
- 6) to take into account the spatial dimension of the analysed phenomenon, which should be presented in the form of maps and plans;
- 7) to urge the authors of academic texts to make public their data sets.

**Keywords:** quantitative methods, data presentation