

# Analiza ilości małżeństw i bezrobocia kobiet w poszczególnych powiatach

Cezary Pukownik, Adam Cencek

## Przygotowanie danych

Dane do analizy zostały pobrane ze strony Głównego Urzędu Statystycznego (GUS) z Banku Danych Lokalnych. Użyto współczynnik feminizacji, liczbę ludności, liczbę małżeństw na 1000 osób oraz liczbę bezrobotnych w poszczególnych powiatach. Dane dla roku 2018.

```
feminizacja = readxl::read_excel('../data/Feminizacja_Gminy.xlsx', sheet = 'TABELICA')
ludnosc = readxl::read_excel('../data/Ludnosc.xlsx', sheet = 'TABELICA')
malzenstwa = readxl::read_excel('../data/Malzenstwa_Gminy.xlsx', sheet = 'TABELICA')
bezrobocie = readxl::read_excel('../data/Bezrobocie_rejestrowne.xlsx', sheet = 'TABELICA')
```

```
kable(head(feminizacja))
```

Kod	Nazwa	ogółem
NA	NA	2018
NA	NA	[osoba]
0201000	Powiat bolesławiecki	106
0202000	Powiat dzierżoniowski	110
0203000	Powiat głogowski	105
0204000	Powiat górowski	102

```
kable(head(ludnosc))
```

Kod	ludnosc	ludn.m	ludn.k
NA	ogółem	mężczyźni	kobiety
NA	2018	2018	2018
NA	[osoba]	[osoba]	[osoba]
0201000	90200	43798	46402
0202000	101437	48355	53082
0203000	89541	43598	45943

```
kable(head(malzenstwa))
```

Kod	Nazwa	ogółem	... 4
NA	NA	2018	2019

Kod	Nazwa	ogółem	... 4
NA	NA	[ - ]	[ - ]
0201000	Powiat bolesławiecki	419	-
0202000	Powiat dzierżoniowski	392	-
0203000	Powiat głogowski	402	-
0204000	Powiat górowski	176	-

```
kable(head(bezrobocie))
```

Kod	Nazwa	mężczyźni	kobiety
NA	NA	2018	2018
NA	NA	[osoba]	[osoba]
0201000	Powiat bolesławiecki	519	764
0202000	Powiat dzierżoniowski	840	914
0203000	Powiat głogowski	908	1694
0204000	Powiat górowski	664	1007

Całość danych po obróbce i sparowaniu z regionami można zestawić w jednej wielokolumnowej tabeli.

```
## # A tibble: 6 x 11
##   Kod Powiat Województwo feminizacja małżeństwa bezrobocie.m bezrobocie.k
##   <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl>
## 1 1815~ powiat podkarpacki 102     447    1474   1952
## 2 1410~ powiat mazowiecki 101     147    405     474
## 3 1418~ powiat mazowiecki 109     818    1806   1681
## 4 1425~ powiat mazowiecki 101     835    4687   4802
## 5 1427~ powiat mazowiecki 103     268    1317   1789
## 6 1430~ powiat mazowiecki 100     187    1752   1789
## # ... with 4 more variables: ludność <dbl>, ludn.m <dbl>, ludn.k <dbl>,
## #   geometry <MULTIPOLYGON [°]>
```

## Analiza - czy liczba małżeństw ma wpływ na bezrobocie u kobiet?

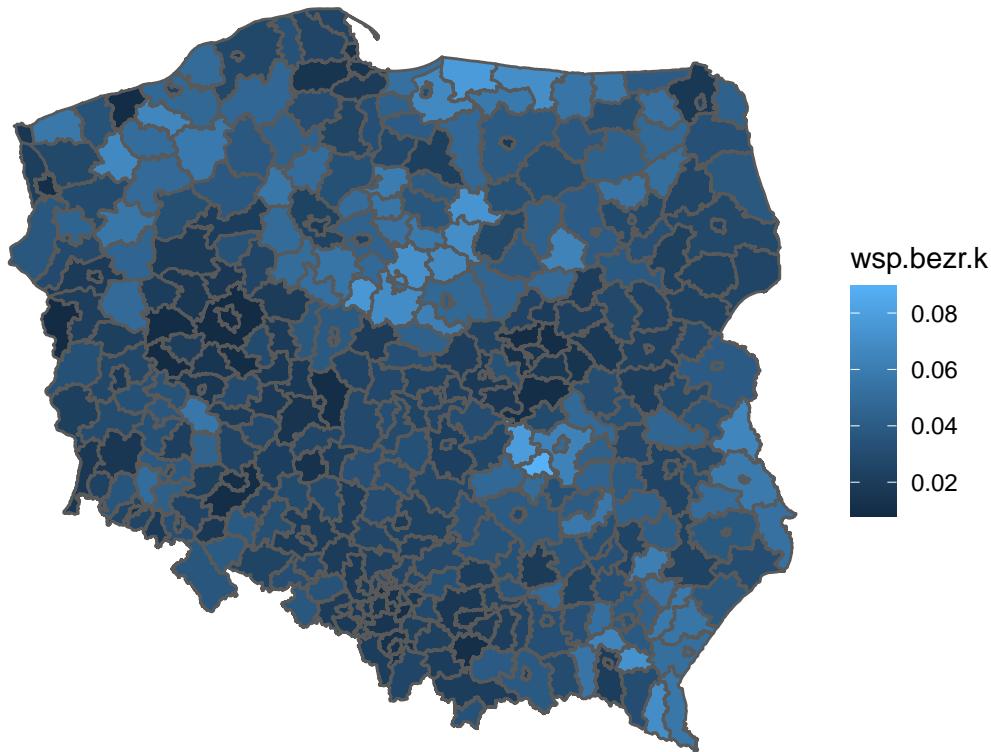
Aby wykluczyć wpływ wielkości ludności na bezrobocie, obliczono współczynnik małżeństw i bezrobocia, jako stosunek miary do ludności.

Powiat	wsp.malz	wsp.bezr.k	wsp.bezr.m
powiat ropczycko-sędziszowski	0.0060102	0.0518597	0.0401263
powiat łosicki	0.0047384	0.0304217	0.0262272
powiat piaseczyński	0.0044228	0.0174526	0.0203762
powiat radomski	0.0054892	0.0629086	0.0618468
powiat sierpecki	0.0051345	0.0675298	0.0512372

Powiat	wsp.malz	wsp.bezr.k	wsp.bezr.m
powiat szydłowiecki	0.0046909	0.0895575	0.0880933

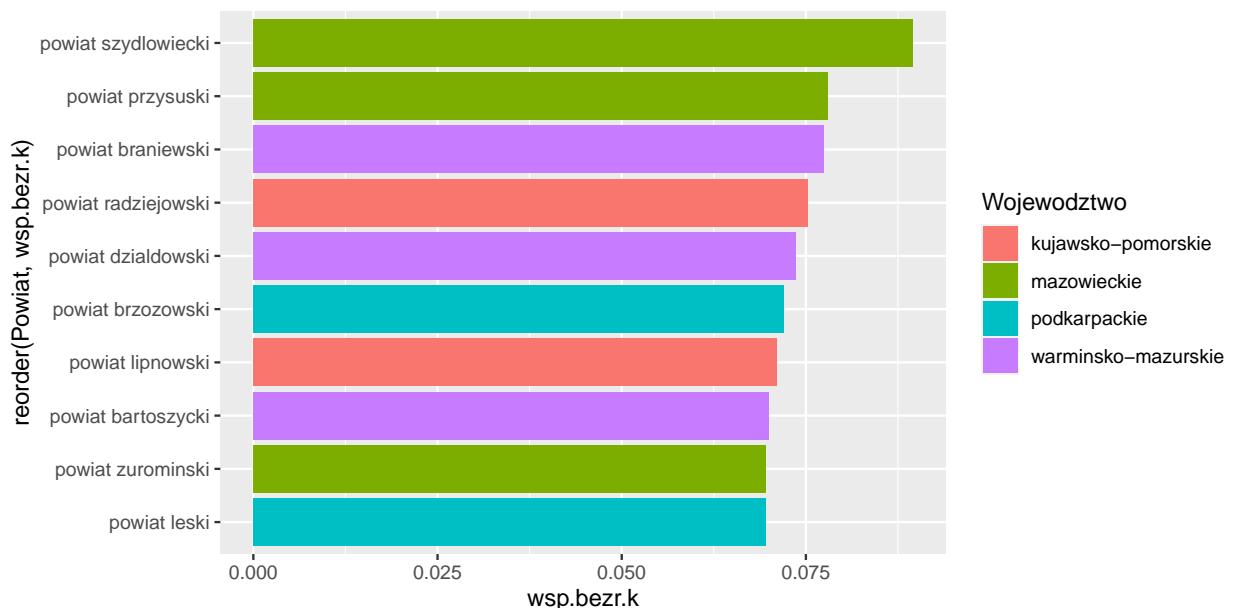
Obliczenia można przedstawić jako mapa ciepła Polski w podziale na powiaty

Współczynnik procentowy bezrobocia kobiet w Polsce

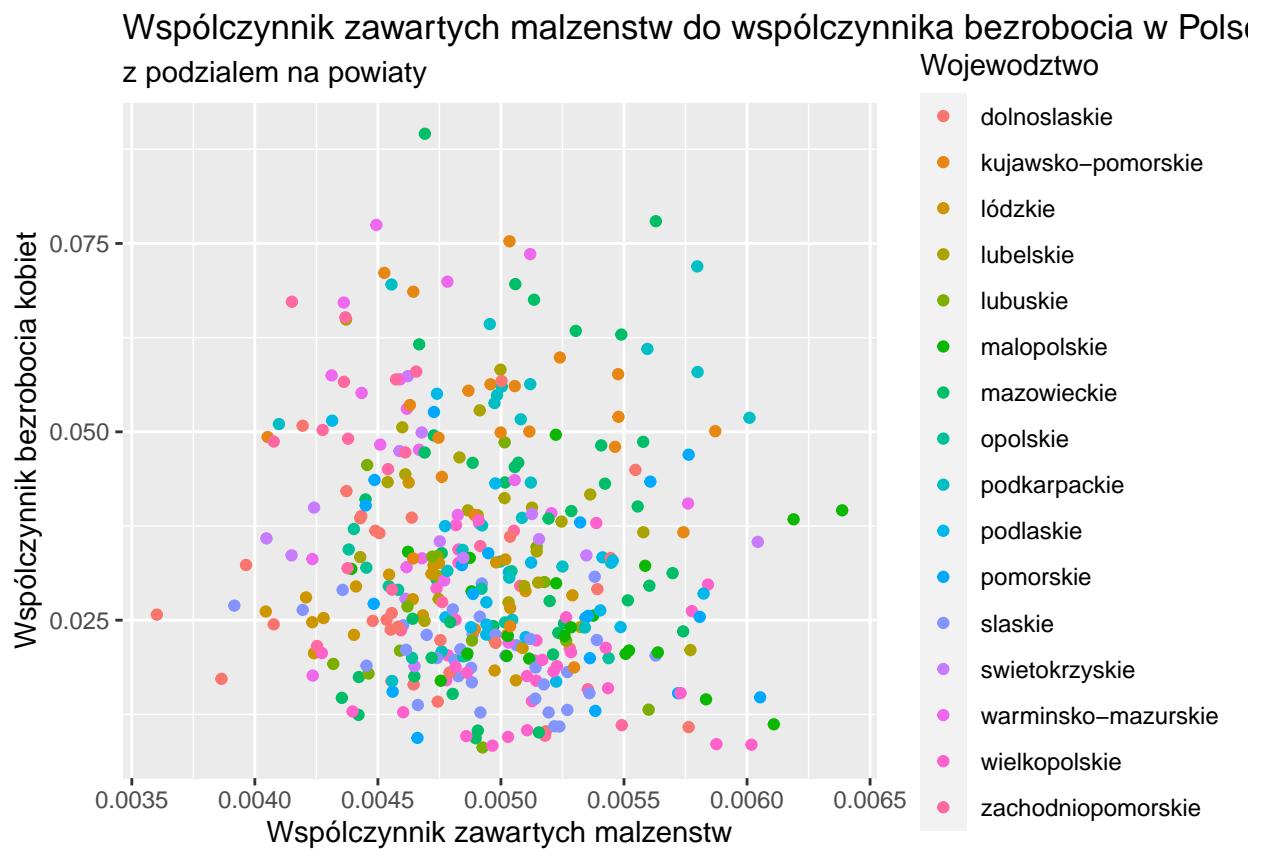


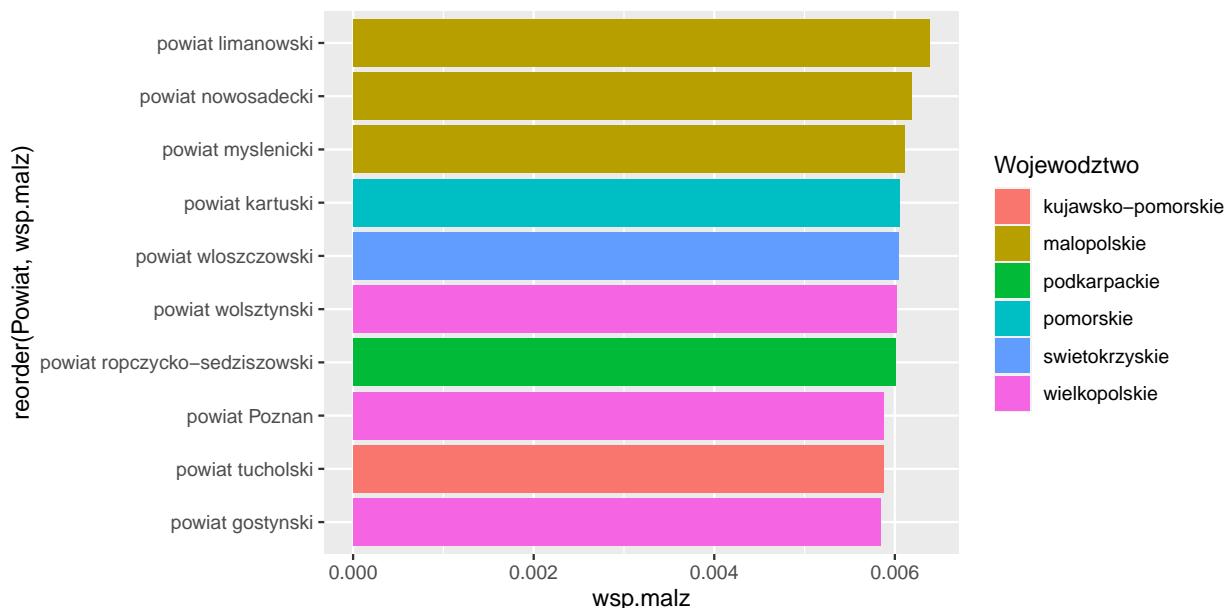
Największy współczynnik bezrobocia kobiet, jest w powiecie szydłowieckim w woj. mazowieckim i wynosi ponad 8%

Powiat	Województwo	wsp.bezr.k
powiat szydłowiecki	mazowieckie	0.0895575
powiat przysuski	mazowieckie	0.0779516
powiat braniewski	warmińsko-mazurskie	0.0774617
powiat radziejowski	kujawsko-pomorskie	0.0752892
powiat działdowski	warmińsko-mazurskie	0.0735856
powiat brzozowski	podkarpackie	0.0719557
powiat lipnowski	kujawsko-pomorskie	0.0710930
powiat bartoszycki	warmińsko-mazurskie	0.0699258
powiat żuromiński	mazowieckie	0.0696000
powiat leski	podkarpackie	0.0695529



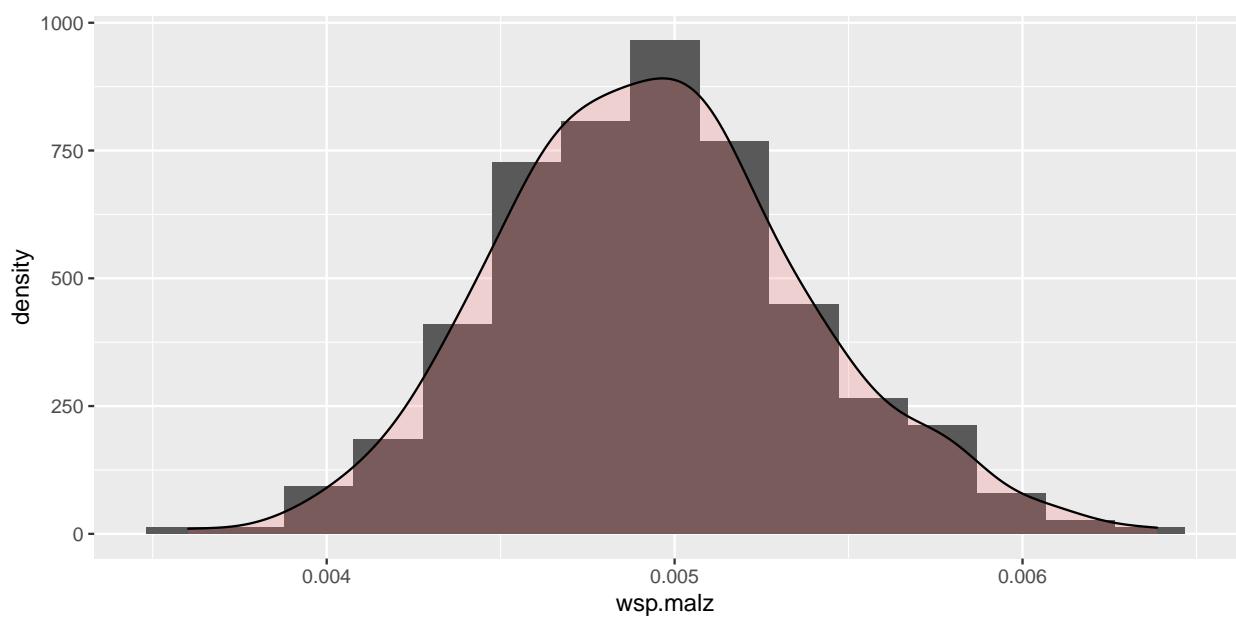
Czy liczba małżeństw jest skorelowana z bezrobociem kobiet?

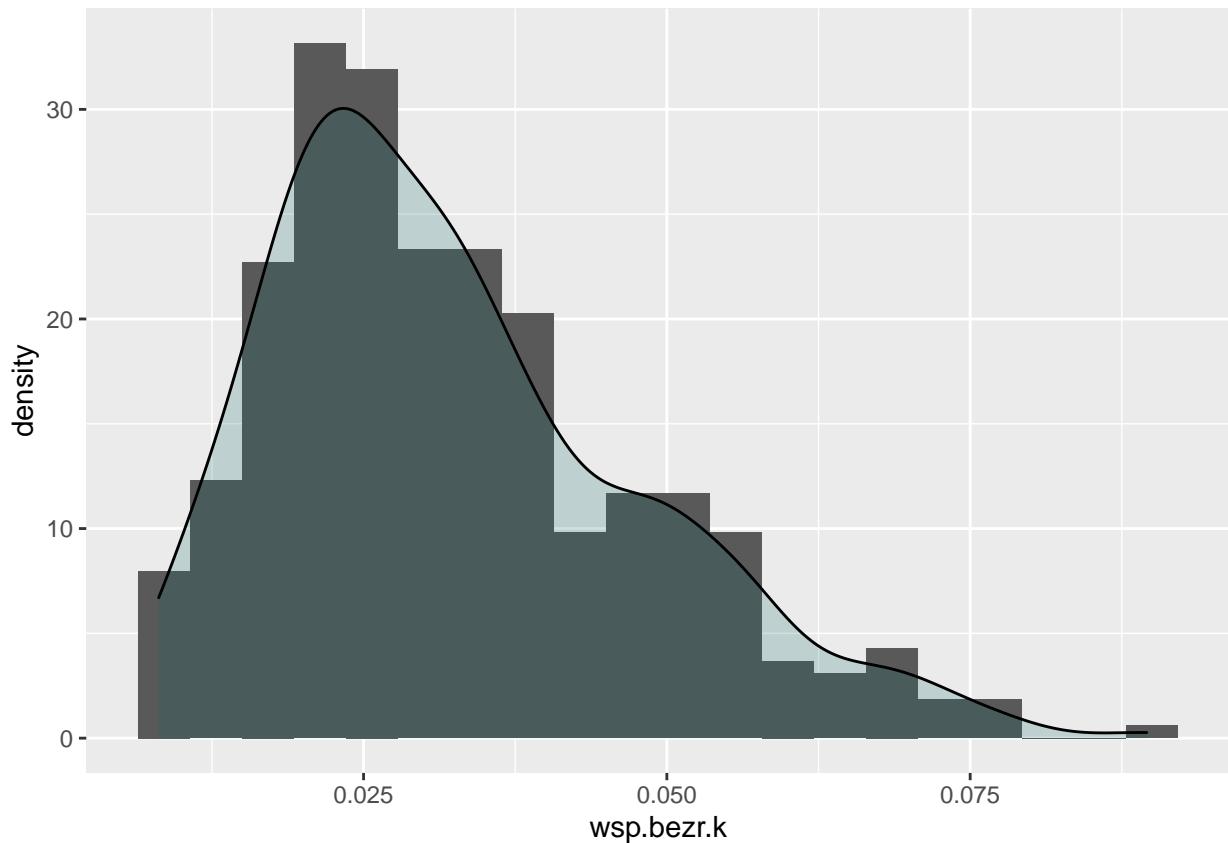




Na wykresie rozrzutu nie widać korelacji między tymi dwoma zmiennymi. Wygląda na to, że liczba małżeństw, nie ma wpływu na bezrobocie kobiet w Polsce.

Można jeszcze obliczyć współczynnik korelacji Pearsona. Jest to jednak test parametryczny zakładający rozkład zbliżony do normalnego badanych zmiennych i warto się upewnić, że te wymaganie jest spełnione

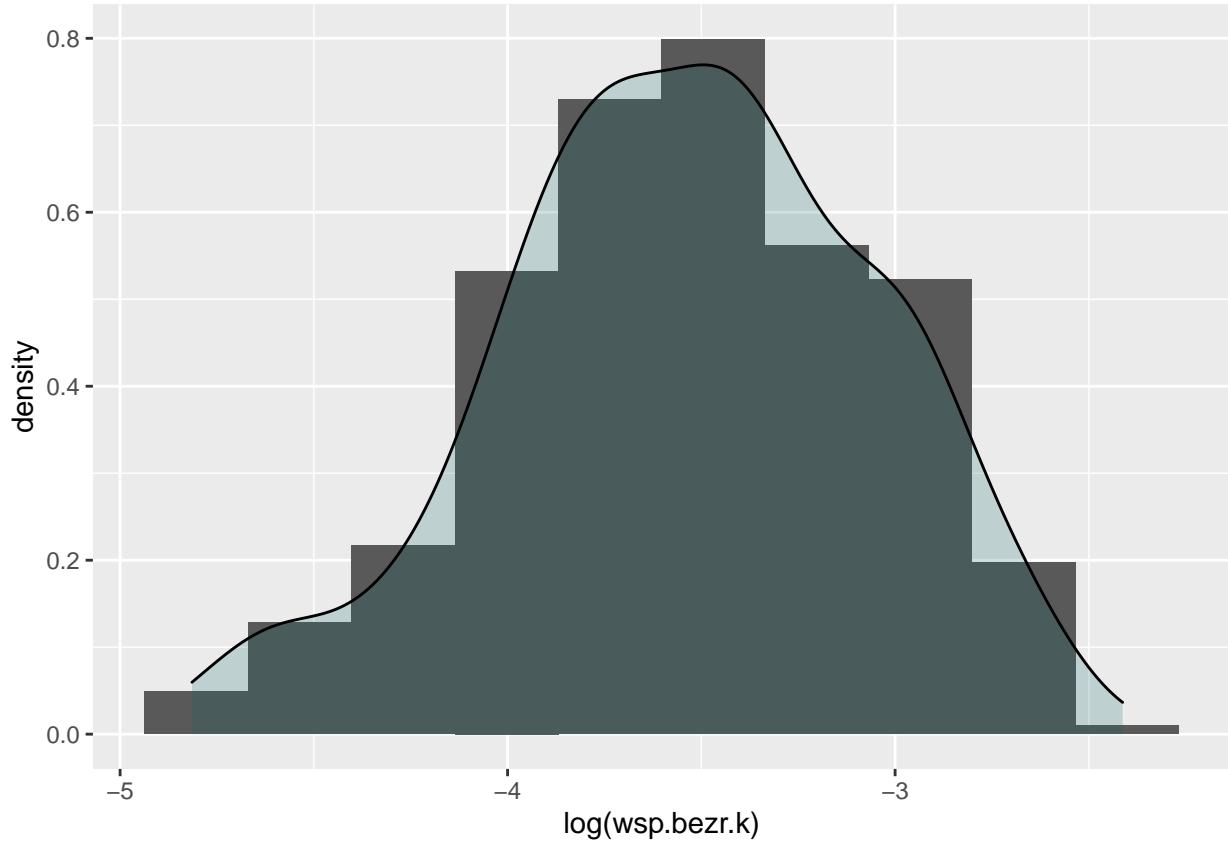




widać, że współczynnik bezrobocia wśród kobiet nie przyjmuje zadowalającego kształtu, różnica nie jest duża aczkolwiek bezpieczniej będzie obliczyć nieparametryczny test korelacyjny Spearmana

```
##
##  Spearman's rank correlation rho
##
## data: df$wsp.malz and df$wsp.bezr.k
## S = 10305254, p-value = 0.01339
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## -0.1268398
```

co powinno być wynikiem zbliżonym do współczynnika Pearsona dla zlogarytmowanej wartości współczynnika bezrobocia wśród kobiet



```
##
## Pearson's product-moment correlation
##
## data: df$wsp.malz and log(df$wsp.bezr.k)
## t = -2.4848, df = 378, p-value = 0.0134
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.22451049 -0.02650822
## sample estimates:
##       cor
## -0.126772
```

*Dla poównania współczynnik Pearsona dla niezlogarytmizowanych danych*

```
##
## Pearson's product-moment correlation
##
## data: df$wsp.malz and df$wsp.bezr.k
## t = -1.9739, df = 378, p-value = 0.04912
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.1995809771 -0.0004095483
## sample estimates:
##       cor
```

## -0.1010072

*wykres rozrzutu i współczynnik Spearmana wykluczają jednak korelację*