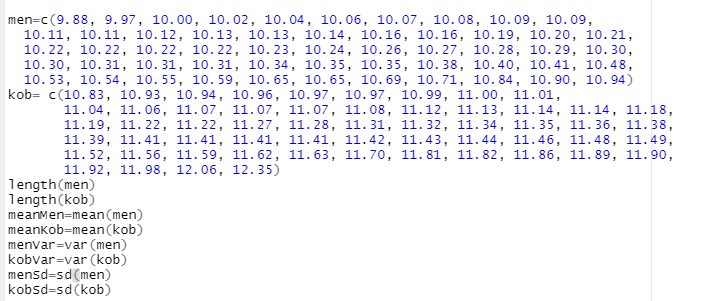
**Praca domowa nr 2**

**Mateusz Markowski gr 1.2/4**

**Zadanie 1**

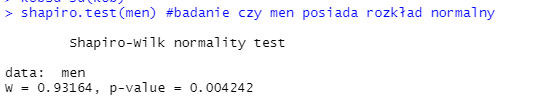
Przygotowanie danych oraz obliczenie dla nich odpowiednio wielkości próbki, średniej wartości próbki, wariancji oraz odchylenia standardowego.

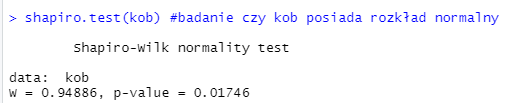
****

Następnym krokiem było sprawdzenie czy dane posiadają rozkład normalny, w tym celu wykorzystano test Shapiro-Wilka.

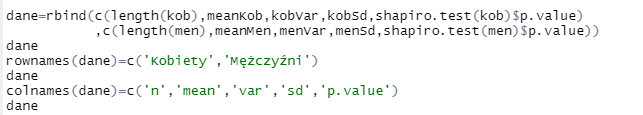


Wyniki dla przeprowadzonych testów:



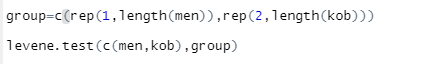


Z otrzymanych wyników widać, że zarówno próbki mężczyzn jak i kobiet nie posiadają rozkładu normalnego, dlatego w celu sprawdzenia hipotezy o równości wariancji nie możemy posłużyć się testem var. Do potwierdzenia bądź odrzucenia hipotezy posłużymy się testem Lavene’a. Można wykorzystać też np.: test Bartletta. Przed przystąpieniem do testu utworzymy jeszcze tabelę wynikową dla danych zawierających m.in. p-value.

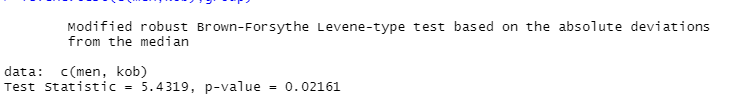




W celu przeprowadzenia testu Levene’a należy pogrupować dane oraz długość odpowiednich wektorów zastąpić dowolną zmienną. W naszym przypadku 1 oznaczono mężczyzn, a 2 kobiety.

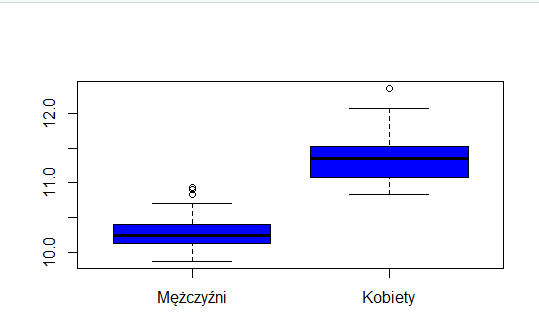


Wyniki testu Levene’a:



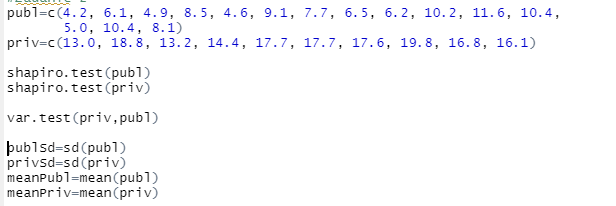
Z otrzymanych wyników że wartość p-value<0.05 więc należy odrzucić hipotezę o równości wariancji w tych populacjach.



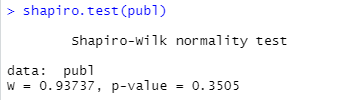


**Zadanie 2**

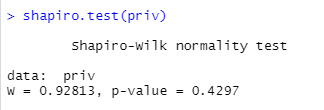
Pierwszym krokiem jest zbadanie czy rozkłady w prywatnych i publicznych koledżach posiadają rozkład normalny. W tym celu wykorzystano test Shapiro-Wilka



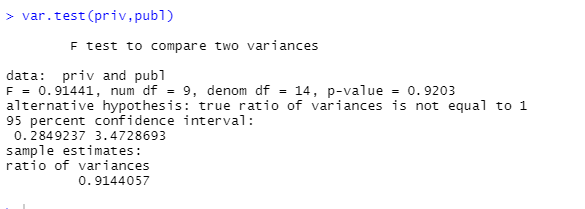
Wyniki testu Shapiro-Wilka dla publicznych danych p-value>0.05 czyli posiada rozkład normalny



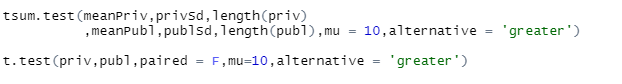
Wyniki testu Shapiro-Wilka dla prywatnych danych p-value>0.05 czyli posiada rozkład normalny.



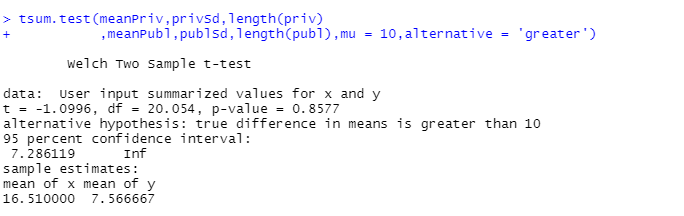
Z dwóch powyższych zrzutów ekranu widać, czemu zdecydowano się na wykonanie testu var.test do analizy wariancji. Wyniki var.test przedstawiono poniżej:

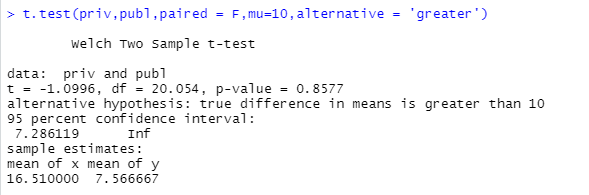


Wartość p-value >0.05 więc możemy potwierdzić hipotezę o równości wariancji. Teraz przechodzimy do badania równości średnich. W tym celu zastosowano tsum.test, do którego potrzebujemy danych dotyczących odchylenia standardowego oraz średnich dla próbek prywatnych i publicznych. Parametry ustawiono mu=10, ponieważ interesuje nas różnica równa lub większa10tys. W tym celu parametr alternative ustawiono na ‘greater.



Wyniki testu:

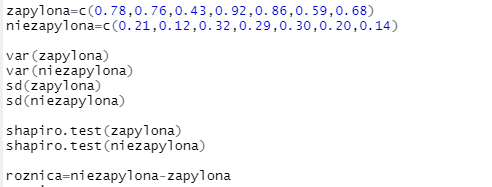




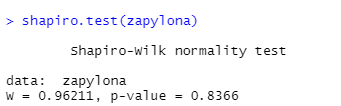
Wartość p-value>0.05 oznacza, że nie można odrzucić hipotezy h0.

**Zadanie 3**

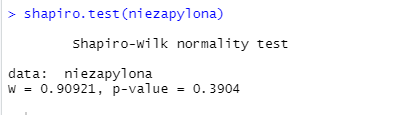
Przygotowanie danych oraz zbadanie czy dane posiadają rozkład normalny. Obliczenie różnicy pomiędzy danymi dotyczącymi zapylenia i nie zapylenia rośliny.



Wyniki testu Shapiro-Wilka dla próbki zapylona przedstawiono poniżej. Jak widać wartość p-value>0.05



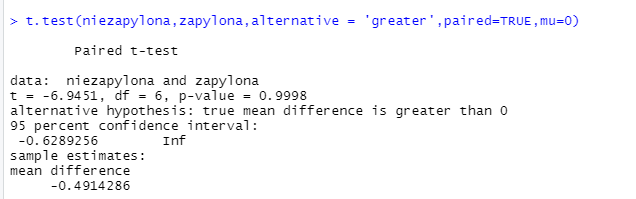
Wyniki testu Shapiro-Wilka dla próbki nie zapylona przedstawiono poniżej. Jak widać wartość p-value>0.05



Z otrzymanych wyników widać, że obie próbki przeszły próbę rozkładu normalnego więc do sprawdzenia badanej hipotezy posłużono się t.testem.

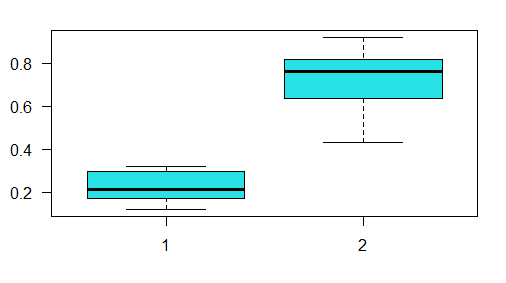


Wyniki przedstawiono poniżej:



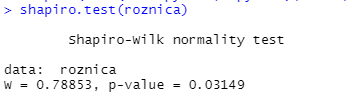
Jak widać wartość p-value>0.05 stąd nie należy odrzucać hipotezy H0.Wyniki dla próbek zapylona oraz niezapylona przedstawiono za pomocą wykresu boxplot:



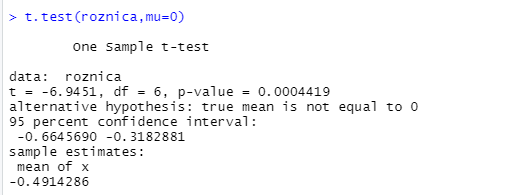


Test Shapiro-Wilka dla danych roznicy=niezapylona-zapylona:



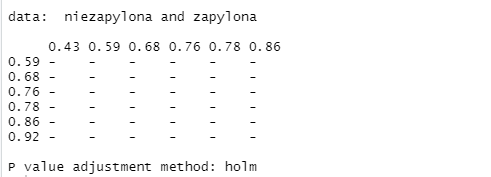


Wartość p-value>0.05. Po przeprowadzeniu t.test-u dla danych roznicy wyniki przedstawiono na poniższym zrzucie ekranowym:



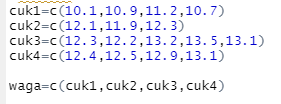
Wyniki dla pairwise.t.test przedstawiono poniżej:



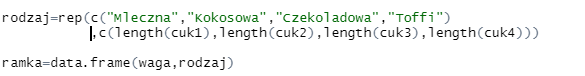


**Zadanie 4**

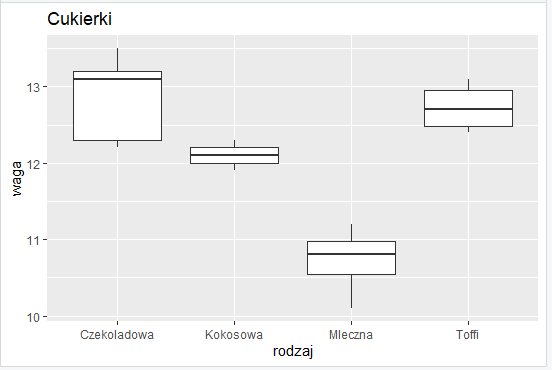
Przygotowanie danych oraz utworzenie wektora z danymi



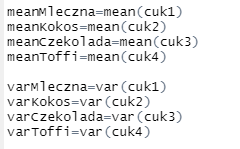
Podpisanie danych odpowiednim rodzajem cukierków oraz utworzenie ramki danych zawierającej wagi i nazwy cukierków. Przedstawienie danych w formie graficznej w postaci qplot.



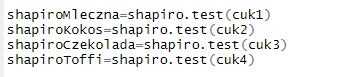




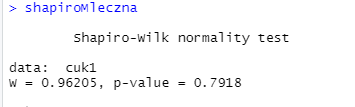
Wykonanie obliczeń średniej oraz wariancji dla każdego rodzaju cukierków:

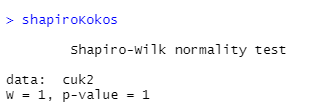


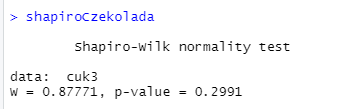
Zbadanie każdego z czterech rodzajów cukierków testem Shapiro-Wilka w celu sprawdzenia czy mamy doczynienia z rozkładem normalnym.

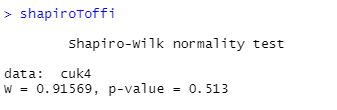


Wyniki testu Shapiro-Wilka:



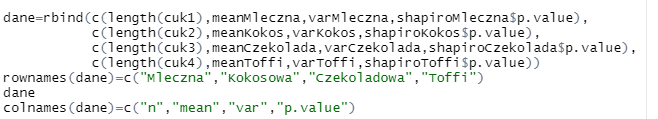


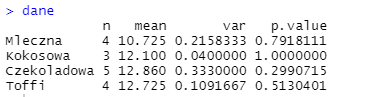




Dla każdego z przedstawionych powyżej testów Shapiro-Wilka widać, że wartość p-value>0.05 co oznacza, że nie możemy odrzucić hipotezy zerowej, że próbki posiadają rozkład normalny.

Przedstawienie wyników w tabeli.



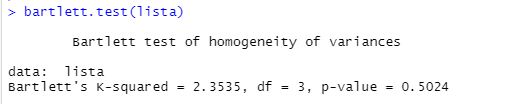


b)

W celu weryfikacji hipotezy o równości wariancji wag wszystkich cukierków należy wykonać test Bartletta. W tym celu grupujemy dane w liste. Następnie na tej liście wykonujemy test Bartletta.



Wyniki testu Bartletta:



Z otrzymanych wyników odczytać możemy, że nie możemy odrzucić hipotezy H0 o równości wariancji wszystkich cukierków, ponieważ wartość p-value>0.05.