

R w analizie statystycznej dla genetyków		
Instytut Genetyki Roślin, Polska Akademia Nauk	prof. dr hab. Idzi Siatkowski idzi@up.poznan.pl Koordynator	Prowadzący: naukowcy z Katedry Metod Matematycznych i Statystycznych Uniwersytetu Przyrodniczego w Poznaniu

Informacja ogólna:

Liczba / formy zajęć	20 godzin lekcyjnych (wykłady i ćwiczenia)
Semestr	Wiosna-lato 2019
Język	angielski
Punkty ECTS	2

Cel kursu:

Poznanie podstawowych możliwości obliczeniowych i graficznych stosowanych w statystyce dla genetyków

Zakres tematyczny:

1. Wprowadzenie do R i RStudio
2. Obliczenia w R
3. Przygotowanie danych
4. Wizualizacje
5. Testowanie
6. Badanie zależności cech
7. Regresja liniowa i wielokrotna
8. Zaliczenie

Efekty (wiedza i umiejętności uzyskane podczas kursu)

Doktorant:

- Wie co to jest R i RStudio
- Zna zalety R i RStudio
- Potrafi zainstalować R i RStudio oraz odpowiednie pakiety
- Zna metody szukania pomocy
- Potrafi wykonać obliczenia matematyczne
- Rozumie pojęcie zmiennej, wektora, macierzy oraz ramki danych
- Potrafi przygotować dane do obliczeń
- Umie wczytać ramki danych, dane tekstowe oraz dane z excela
- Potrafi wykonać graficzną prezentację danych
- Umie przedstawić funkcję w postaci wykresu
- Zna reguły stosowania testów statystycznych
- Umie wykonać testy dla wartości średnich
- Potrafi zinterpretować wyniki otrzymane po testowaniu
- Potrafi porównać statystycznie dwie i więcej grup
- Umie przeprowadzić badanie zależności cech
- Rozumie znaczenie regresji liniowej i wielokrotnej
- Potrafi wykonać regresję liniową i wielokrotną
- Umie zinterpretować wyniki otrzymane po zastosowaniu regresji liniowej i wielokrotnej

Zawartość merytoryczna kursu:

- 1. Wprowadzenie do R i RStudio (1 godz.)**
 - 1.1. Co to jest R
 - 1.2. Zalety R
 - 1.3. Instalacja R i RStudio
 - 1.4. Pakiety
 - 1.5. Dokumentacja i szukanie pomocy
- 2. Obliczenia w R (3 godz.)**
 - 2.1. Proste obliczenia matematyczne
 - 2.2. Zmienne
 - 2.3. Wektory liczbowe i znakowe oraz macierze i ramki danych
- 3. Przygotowanie danych (2 godz.)**
 - 3.1. Wczytanie ramki danych
 - 3.2. Wczytanie danych tekstowych
 - 3.3. Wczytanie danych z Excela
 - 3.4. Wczytanie danych różnych formatów
- 4. Wizualizacje (3 godz.)**
 - 4.1. Graficzna prezentacja danych
 - 4.2. Wykresy dla przykładowych funkcji
- 5. Testowanie (4 godz.)**
 - 5.1. Wprowadzenie
 - 5.2. Dwie populacje - rozkłady normalne
 - 5.3. Dwie populacje - dowolne rozkłady
 - 5.4. Analiza wariancji - ANOVA
 - 5.5. Testy wielokrotne
- 6. Badanie zależności cech (1 godz.)**
 - 6.1. Korelacje
 - 6.2. Tablice kontyngencji
- 7. Regresja liniowa i wielokrotna (4 godz.)**
 - 7.1. Regresja liniowa
 - 7.2. Regresja wielokrotna
 - 7.3. Selekcja zmiennych
- 8. Zaliczenie (2 godz.)**

Techniki/metody nauczania:

- wykłady w języku angielskim z wykorzystaniem technik multimedialnych
- ćwiczenia przy komputerach

Ocena efektów uczenia się:

- egzamin praktyczny przy komputerach