

Journée Julia pour les statistiques

Rapport sur les contributions

Fondamentaux

lundi 15 juin 2026 09:00 (1h 30m)

Objectifs : Maîtriser les bases de Julia pour la manipulation de données

- **Installation et environnement** (15 min)
 - Installation de Julia et VS Code/Jupyter
 - Packages essentiels : Pkg.add(["DataFrames", "Statistics", "StatsBase", "Plots", "CSV"])
- **Structures de données** (45 min)
 - Vecteurs, matrices et tableaux multidimensionnels
 - DataFrames : création, indexation, filtrage
 - Manipulation de colonnes et lignes
 - Exercice pratique : charger et explorer un jeu de données CSV
- **Broadcasting et opérations vectorisées** (30 min)
 - Opérateur point (.) pour les opérations élément par élément
 - Fonctions anonymes et map/reduce
 - Exercice : calculs statistiques sur des colonnes

Statistiques descriptives

lundi 15 juin 2026 11:00 (1h 30m)

Objectifs : Calculer et visualiser des statistiques de base

- **Mesures de tendance centrale et de dispersion** (45 min)
 - Moyenne, médiane, mode avec Statistics et StatsBase
 - Variance, écart-type, quantiles
 - Fonction describe() pour un aperçu complet
 - Exercice : analyse descriptive d'un dataset réel
- **Visualisation de données** (60 min)
 - Introduction à Plots.jl et StatsPlots.jl
 - Histogrammes, boxplots, scatter plots
 - Graphiques de distribution
 - Personnalisation des graphiques
 - Exercice : créer un tableau de bord statistique visuel

Distributions et échantillonnage

lundi 15 juin 2026 14:00 (1h 30m)

Objectifs : Travailler avec les distributions de probabilité

- **Package Distributions.jl** (45 min)
- Distributions courantes : Normal, Binomial, Poisson, Uniform
- Générer des échantillons aléatoires
- Calcul de PDF, CDF, quantiles
- Ajustement de distributions à des données
- **Tests statistiques de base** (45 min)
- Tests d'hypothèse avec HypothesisTests.jl
- Test t de Student, test de Wilcoxon
- Test du chi-deux
- Interprétation des p-values
- Exercice : réaliser plusieurs tests sur vos données

Régression et modélisation

lundi 15 juin 2026 16:00 (1h 30m)

Objectifs : Construire des modèles statistiques

- **Régression linéaire avec GLM.jl** (60 min)
- Formules de modèle (@formula)
- Ajustement d'un modèle linéaire simple et multiple
- Interprétation des coefficients et statistiques
- Diagnostic du modèle (résidus, R²)
- Prédictions
- **Projet final** (45 min)
- Analyse complète d'un jeu de données :
 1. Import et nettoyage
 2. Statistiques descriptives
 3. Visualisations
 4. Tests d'hypothèse
 5. Modèle de régression
 6. Interprétation et conclusions