

# SAé 4.02

## Reporting d'une analyse multivariée

Caroline BAZZOLI  
caroline.bazzoli@univ-grenoble-alpes.fr

23 janvier 2024

# Table des matières

## 1 Introduction

- Extraction du PN 2022
- Objectifs
- Organisation générale

## 2 Suivi et évaluation

## 3 Rendu

## 4 Sujet : données et problématiques

## 5 Les attendus

- Compétences ciblées
  - Traiter des données à des fins décisionnelles
  - Analyser statistiquement les données
  - Valoriser une production dans un contexte professionnel
- Apprentissages critiques : *cf grille évaluation*
- Ressources mobilisées
  - R4.01 | Automatisation et test en programmation
  - R4.02 | Méthodes factorielles
  - R4.03 | Classification automatique

- L'étudiant est mis en situation de conception d'un outil de restitution automatisée des résultats d'une analyse exploratoire de données.
  - En regard d'une **problématique métier** présentée à l'étudiant et d'un **entrepôt de données** mis à sa disposition :
    - **Mettre en œuvre un reporting** et en assurer l'**automatisation** pour répondre au caractère récurrent des analyses à effectuer.
    - **Savoir appliquer la démarche d'analyse exploratoire complète** à des données réelles en s'appuyant sur :
      - des méthodes d'analyse statistique descriptives
      - une **première approche du data mining**.
- ⇒ **Valoriser les résultats** ou les **données** au travers d'une mise en récit amenant à une **prise de décision motivée**.

- **Travail équipe** : quadrinômes ou trinôme, choisi librement au sein de votre groupe
  - 46 étudiants
  - 10 groupes de 4 + 1 groupe de 6 (ou 2 groupes de 3)
  - 20 EMS (5 groupes) + 27 VCOD (6 groupes)
- **24h** de cours + **45h** de projet en salle info (voir ADE)
  - **TP Analyse factorielles** en R réalisés par Chantal Bédecarrax (10h)
  - **Reporting/Shiny** réalisées par Olivier Renault (12h)
- Début : semaine du 22 janvier 2024
- Fin : Semaine du 25 mars (semaine bloquée)

- Chaque groupe est tutoré par un-e enseignant-e de statistique
  - Trois rendez-vous de point à fixer avec le tuteur ou la tutrice
  - Évaluation sur 23 items précisant les 10 apprentissages critiques
    - **Compétence 1** : Analyse statistique
      - Note sur la qualité d'une analyse présentée
      - Modulation possible si plusieurs analyses de qualité
    - **Compétence 2** : application R shiny
      - Note sur le guide et la mise en œuvre technique de l'application Rshiny
    - **Compétence 3** : qualité du support de présentation
      - Note sur le visuel / oral de soutenance et démo de l'application
- cf détails dans la grille d'évaluation

- **Petite application (Shiny)** pour présenter des résultats d'analyse multivariée
- **Soutenance le vendredi 29 mars 2024** (après 1 semaine de projet)
  - Synthèse résultats : 10min
  - Démo de l'application : 5min
  - Questions : 5min
- **Guide d'utilisateur/ guide technique** (format libre) et **support de présentation** à déposer au plus tard le **jeudi 28 mars à 12h** sur CHAMILO
- Pas de rapport écrit



- **Descriptif du fichier** : Le fichier fourni contient des données athlétiques sur les joueurs de l'équipe sur l'ensemble des match de ligue 2 de la journée 1 à la journée 17.
  - **Problématique** : identifier des groupes de joueurs en fonction des différentes variables puis de les comparer aux performances des équipes de la même ligue et de la ligue 1.
- ⇒ **Tuteur** : Mme BAZZOLI
- ⇒ **Nombre de groupe** : 3 ou 4 groupes (VCOD/EMS)





- **Descriptif du fichier** : Le fichier fourni contient les informations sur les sorties de randonnée à ski, postées par les contributeurs sur le forum communautaire Skitour pour le massif alpin Vanoise sur les 25 dernières années.

- **Problématique** :

- 1 Analyser la fréquentation du massif Vanoise selon une approche statistique spatio-temporelle pour traiter les questionnements suivants : Où vont préférentiellement les pratiquants ? Quand ? Quelle est la période de pratique d'un sommet ?
- 2 Interpréter les résultats issus des analyses statistiques au regard d'éléments contextuels tels que : jour de semaine/weekend et jours fériés, conditions météorologiques et niveau de risque avalanche.

⇒ **Tuteur** : Mr GARAT

⇒ **Nombre de groupe** : 3 ou 4 groupes (VCOD/EMS)



- **Descriptif du fichier** : Le fichier fourni contient les réponses d'un questionnaire diffusé auprès des étudiant-es sur le campus de l'UGA (par les étudiants du groupe EMS). L'objectif principal de l'enquête était de cibler les étudiants ne connaissant pas l'EOIP (Espace de l'Orientation et de l'Insertion Professionnelle) et de comprendre cette méconnaissance.
- **Problématique** : L'objectif de la SAé est d'identifier les profils des répondants au questionnaire sur la connaissance du service d'insertion et d'orientation professionnelle (OIP) de l'UGA.
  - Profils des étudiants qui ne connaissent pas le service OIP
  - Profils des étudiants qui connaissent le service OIP

⇒ **Tuteur** : Mme BEDECARRAX

⇒ **Nombre de groupe** : 3 ou 4 groupes (VCOD/EMS)

# Données RSE auprès de l'EUT



- **Descriptif du fichier** : Le fichier fourni contient les réponses à un questionnaire réalisé auprès des étudiants de la journée RSE 2023 (connaissance du concept de RSE, Votre place vis à vis du développement durable, Appréciation de la mise en œuvre de la politique de RSE au sein de l'EUT)
  - **Problématique** : Identifier des profils d'étudiants selon leur connaissance ou leur appréciation de mise en œuvre de la politique RSE au sein de l'EUT.
- ⇒ **Tuteur** : Mme BAZZOLI
- ⇒ **Nombre de groupe** : 3 ou 4 groupes (VCOD/EMS)



- **Descriptif du fichier** : Le fichier fourni contient les réponses d'un questionnaire diffusé en ligne auprès de 16 154 moniteurs de ski.

- **Problématique** : : Ce projet de recherche a été soutenu par le SNMSF et financé par le Labex Ittem. Il a été mené par des chercheurs et chercheuses issu-es de trois laboratoires de l'Université Grenoble- Alpes. L'objectif de l'enquête est d'observer les adaptations au changement climatique des moniteurs de ski ESF.
  - Questionnaire organisé en thèmes
  - Choisir au moins deux thèmes par groupe

⇒ **Tuteur** : Mme BAZZOLI

⇒ **Nombre de groupe** : 3 ou 4 groupes (VCOD/EMS)

- Analyses statistiques pour répondre à une problématique
  - Choix raisonné des variables
  - Analyses univariées
  - Analyses multivariées
- ⇒ Méthodes factorielles et approches de classification automatique
- Visualisation interactive des résultats des analyses à l'aide d'une application R Shiny exclusivement sur les données de la SAé
  - **Ne pas faire produire une application généraliste !**
  - Paramètres sur lesquels jouer :
    - Individus/ variables supplémentaires
    - Nombre de classes
    - Choix des graphiques à afficher

Des questions ?