

STT-1100 — Feuille aide-mémoire

Université Laval

```
knitr::opts_chunk$set(echo = TRUE, eval = FALSE, message = FALSE, warning = FALSE)
```

ENVIRONNEMENT DE TRAVAIL

RStudio

- Script: Ctrl + Shift + N
- Rendu Quarto: Ctrl + Shift + K

Quarto

```
---  
title: "Titre"  
format: html  
---
```

Chunk R :

```
::: {.cell}  
```{r .cell-code}  
summary(df)
```:  
:::
```

Packages de base

```
library(tidyverse)
library(readr)
library(UlavalSSD)
library(palmerpenguins)
```

MANIPULER DES DONNÉES — dplyr

Base

```
select(df, col1)
filter(df, col1 > 10)
```

Transformation

```
mutate(df, z = x/y)
arrange(df, desc(x))
```

Regrouper et résumer

```
group_by(df, cat)
summarise(df, moy = mean(x, na.rm = TRUE))
```

VISUALISATION — ggplot2

Numérique

```
ggplot(df, aes(x, y)) + geom_point()
ggplot(df, aes(x)) + geom_histogram()
```

Catégorique

```
ggplot(df, aes(x, fill = cat)) + geom_bar()
ggplot(df, aes(cat, num)) + geom_boxplot()
```

Utiliser `facet_wrap()` pour comparer plusieurs groupes

ANALYSE DESCRIPTIVE

Statistiques numériques

```
mean(x, na.rm=TRUE)
sd(x)
summary(df)
```

Fréquences

```
table(df$cat)
prop.table(table(df$cat))
```

skimr

```
library(skimr)
skim(df)
```

ASTUCES BONNES PRATIQUES

- `glimpse(df)` : aperçu structure
- `head(df)` / `tail(df)`
- Commentaires # pour clarifier
- Commits fréquents dans GitHub
- Éviter les noms de variables ambigus