Mandat dunin

simon-pierre Beugré

2025-05-12

Table of contents

Préface **4**
Présentation du projet 4
Structure du document 4

I Contexte et projet **5**

1 Introduction **6**
1.1 Objectifs du projet 6
1.2 Méthodologie 6
1.3 Équipe et parties prenantes 6
1.4 Durée (6 semaines) 6

II Analyse exploratoire (EDA) **7**

2 Analyse exploratoire des données (EDA) **8**
2.1 Introduction 8
2.2 Structure du chapitre 8

3 Colllecte et Préparation des données **9**
3.1 Sources et méthodes de collecte des données 9
3.2 Nettoyage et prétraitement des données 9
3.2.1 table commandes 10
3.2.2 table contrat 12
3.2.3 table projet 13
3.3 Transformation et enrichissement des données 14
3.3.1 structure de fichier xml 14
3.3.2 extractions des information et tables potentielles des différents fichiers xmls 20
3.3.3 diagrammmes de sélection des entités pour l'analyse 22

4 Analyses descriptives **23**
4.1 Introduction 23
4.2 Méthodologie 23
4.3 Structure de cette section 23

5 Analyse des documents (commandes) 24
5.1 Vue d'ensemble des commandes et préparations des données 24
5.1.1 Résumé de la table 24
5.1.2 selections des variables (colonnes pertinentes) 29
5.1.3 Conversion des colones dans des types de variables appropriés 31
5.1.4 Gestion des valeurs manquantes (NA) 33
5.2 Distribution des commandes par montant 38
5.3 Évolution temporelle des commandes 38
5.4 Commandes par canal 38
5.5 Commandes par région 38
5.6 Taux de conversion 38
5.7 Synthèse de l'analyse des commandes 38

III Analyse avancée 39

6 Titre à remplir 40

7 À compléter 41

IV Application finale 42

8 Titre à remplir 43

9 À compléter 44

Préface

Bienvenue dans le projet “Mandat dunin”. Ce document présente l’ensemble du travail réalisé pour analyser les données, créer une architecture de données robuste, et développer une application finale.

Présentation du projet

Ce projet vise à mettre en place une solution d’analyse automatisée des données de vente afin d’identifier les tendances clés et les opportunités d’amélioration de la performance commerciale..

Structure du document

Ce document est organisé en plusieurs parties:

1. **Contexte et projet**: Présentation du contexte, des objectifs et de la méthodologie.
2. **Collecte et Analyse exploratoire (EDA)** des donnés :Statistiques descriptives, visualisations et analyse des relations entre les variables.
3. **Analyse avancées**: Techniques d’analyse avancées appliquées aux données.
4. **Mise en oeuvre des tableaux de bords** : Présentation de l’application développée à partir des analyses.

Part I

Contexte et projet

1 Introduction

Ce document introduit le projet **Mandat Dunin**. Il s'agit d'un mandat d'analyse de données visant à extraire de la valeur à partir des données commerciales et opérationnelles d'une entreprise, en vue d'optimiser ses performances de ventes et marketing.

1.1 Objectifs du projet

Mettre en place une solution d'analyse automatisée des données de vente afin d'identifier les tendances clés et les opportunités d'amélioration de la performance commerciale.

1.2 Méthodologie

la méthodologie de gestion de projet choisie pour ce projet est le : scrum .

1.3 Équipe et parties prenantes

Datcha simon-pierre Beugré : Consultant

Marc André : Encadreur , product owner

1.4 Durée (6 semaines)

du 15 mars au 29 avril 2025

Part II

Analyse exploratoire (EDA)

2 Analyse exploratoire des données (EDA)

2.1 Introduction

Ce chapitre présente l'analyse exploratoire des données réalisée dans le cadre du projet. Il comprend des statistiques descriptives, des visualisations et une analyse des relations entre les variables.

2.2 Structure du chapitre

Ce chapitre est organisé en plusieurs sections:

1. **Collecte et Préparation des données**
2. **Analyse descriptive** : Analyse détaillée des commandes, contrats et kits , clients.
3. **Visualisations**: Représentations graphiques des données pour identifier les tendances et les patterns.

3 Colllecte et Préparation des données

3.1 Sources et méthodes de collecte des données

Table 3.1: Vue d'ensemble des sources de données

Source	Type	Description	Methode_de_collecte	Volumétrie	Fréquence
avivia database	Base de données SQL	Base de donnée transactionnel des activités de ventes	Via la plateforme Eve	8 Go	Temps réel

3.2 Nettoyage et prétraitement des données

connexion et vue d'ensemble de la base de données

```
library(DBI)
library(odbc)

conn <- dbConnect(odbc(), .connection_string = "Driver={ODBC Driver 18 for SQL Server};",
            Server = "server-sql-dunin-eve.database.windows.net", Database = "AviviaEve_Copy"
            UID = "accesAviviaData", PWD = "Citric&Private&Designed&Maturity1&Cloning")


# selectionner les noms des tables
print(DBI::dbListTables(conn,schema="dbo"))
```

```
 [1] "CommEW"          "CommEWArch"      "CommHistoEW"     "CommLogEW"
 [5] "ComposRepriseEW" "ContratEW"       "ContratHistoEW"  "ContratLogEW"
 [9] "DeductionEW"     "eCommerceEW"     "EventEW"         "exportProjet"
[13] "ProjetEW"        "ProjetLogEW"     "SequenceEW"      "SessionEW"
[17] "SessionExtEW"    "Sessions"        "sysdiagrams"     "UsagerEW"
[21] "UsagerInfoEW"    "UsagerLogEW"     "UsagerParamRole"
```

les informations de la base de données sont principalement contenues dans trois tables principales : ProjetEW, CommEW et ContratEW.

```
library(dplyr)
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
#acces aux tables  sans chargement en memoire
GetLazyDBAccess <- function(conn,tables){
  lazyDb <- lapply(tables, function(tbl_name) {
    tbl(conn, tbl_name) # Accesses each table from the connection
  })
  lazyDb
}

mainTables <- c("CommEW","ContratEW","ProjetEW")
AviviaDBData <- GetLazyDBAccess(conn,mainTables)
names(AviviaDBData) <- mainTables



print(AviviaDBData |> names())
```

```
[1] "CommEW"    "ContratEW" "ProjetEW"
```

3.2.1 table commandes

```
AviviaDBData$CommEW |> glimpse()
```

```
Rows: ??
Columns: 37
Database: Microsoft SQL Server 12.00.0429[accesAviviaData@server-sql-dunin-eve/AviviaEve_Copy
$ Code                <chr> "C-000089", "C-000093", "C-000115", "C-000119", "C~
$ Projet              <chr> "P-000175", "P-000181", "P-000219", "P-000223", "P~
$ Archive             <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
$ Abandon             <int> 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0,~
$ Emplacement         <chr> "Manufacturier", "Manufacturier", "Client", "Clien~
$ Statut              <chr> "Installation", "Installation", "Installation", "I~
$ Createur            <chr> "dunintech", "glebel", "clientweb", "clientweb", "~
$ Editeur             <chr> "", "", "", "", "", "", "", "", "", "", "", "", ""~
$ DocumXML            <chr> "<xmlDunin version='1.00'><obj clas='Comm'><prop n~
$ EveXML              <chr> "<xmlDunin version='1.00'><obj clas='Ingenierie'><~
$ ERPXML              <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ FlagERP             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ DocumRef            <chr> "", "", "", "", "", "", "", "", "", "", "", "", ""~
$ Avis                <int> 0, 0, 1, 1, 1, 0, NA, 0, 1, 1, 1, NA, 0, 0, 1, 0, ~
$ RequisMatXML        <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ EstimeNestingXML    <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ RequisNestingXML    <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ CtrlNesting         <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ FicCNCSauve         <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ EtapeTransfertERP   <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ Proprietaire        <chr> "dunintech", "glebel", "clientweb", "clientweb", "~
$ BomXML              <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ DateCreation        <dttm> 2023-08-30 14:48:34, 2023-08-31 13:11:14, 2023-09~
$ LienERP             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ EtiquettesHTML      <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ AdresseCourrielComm <chr> "{\"infos\":[{\"etape\":\"Saisie\",\"email_From\":~
$ TypeProcess         <chr> "interne", "interne", "externe", "externe", "exter~
$ bomJSON             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ export1             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ export2             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ export3             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ export4             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ export5             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ refExterne          <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ timeStamp           <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ ref1                <chr> "", "", "", "", "", "", "", "cuisine 2", "", "test~
$ ref2                <chr> "", "", "", "", "", "", "", "", "", "", "", "", ""~
```

3.2.2 table contrat

```
AviviaDBData$ContratEW |> glimpse()
```

```
Rows: ??
Columns: 38
Database: Microsoft SQL Server 12.00.0429[accesAviviaData@server-sql-dunin-eve/AviviaEve_Copy
$ Code                <chr> "P-001666-S1", "P-001471-S8", "P-000907-S1", "P-00~
$ Projet              <chr> "P-001666", "P-001471", "P-000907", "P-001681", "P~
$ Archive             <int> 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0,~
$ Abandon             <int> 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,~
$ Emplacement         <chr> "Manufacturier", "Manufacturier", "Manufacturier",~
$ Statut              <chr> "Saisie", "Saisie", "Saisie", "Saisie", "Saisie", ~
$ Createur            <chr> "alaplante", "abreault", "abreault", "rbouchard@av~
$ Editeur             <chr> "", "", "", "", "", "", "", "", "", "", "", "", ""~
$ DocumXML            <chr> "<xmlDunin version='1.00'><obj clas='Contrat'><pro~
$ EveXML              <chr> "<xmlDunin version='1.00'><obj clas='Ingenierie'><~
$ ERPXML              <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ DocumRef            <chr> "", "", "", "", "", "", "", "", "", "", "", "", ""~
$ DernierJourValide   <date> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ Avis                <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
$ RequisMatXML        <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ EstimeNestingXML    <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ RequisNestingXML    <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ CtrlNesting         <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ FicCNCSauve         <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ EtapeTransfertERP   <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ Proprietaire        <chr> "alaplante", "abreault", "abreault", "rbouchard@av~
$ BomXML              <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ DateCreation        <dttm> 2025-01-09 21:23:06, 2025-01-13 15:13:06, 2025-02~
$ LienERP             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ EtiquettesHTML      <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ FlagERP             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ AdresseCourrielComm <chr> "{\"infos\":[{\"etape\":\"Saisie\",\"email_From\":~
$ TypeProcess         <chr> "interne", "interne", "interne", "interne", "inter~
$ bomJSON             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ export1             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ export2             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ export3             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ export4             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ export5             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
```

```
$ refExterne          <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ timeStamp           <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ ref1                <chr> "Cuisine V1", "Plan cuisine (version 2)", "Cuisine~
$ ref2                <chr> "", "", "", "", "", "", "", "", "", "", "", "", ""~
```

3.2.3 table projet

```
AviviaDBData$ProjetEW |> glimpse()
```

```
Rows: ??
Columns: 7
Database: Microsoft SQL Server 12.00.0429[accesAviviaData@server-sql-dunin-eve/AviviaEve_Copy
$ Code         <chr> "P-000001", "P-000002", "P-000004", "P-000006", "P-000011~
$ Client       <chr> "Avivia", "", "", "", "", "Avivia", "", "CRobitaille", "A~
$ Createur     <chr> "1", "1", "glebel", "glebel", "glebel", "1", "glebel", "c~
$ ProjetXML    <chr> "<xmlDunin version='1.00'><obj clas='Projet'><prop nom='$~
$ Archive      <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
$ DateCreation <dttm> 2022-07-21 12:33:39, 2022-10-12 18:18:16, 2022-10-12 20:~
$ Libre        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
```

Après une analyse des champs des différentes tables, nous avons choisi de nous concentrer sur ceux contenant des fichiers XML notemment : ***DocumXML***, ***EveXML*** et *ProjetXML* . Ces derniers non seulement synthétisent les informations de chaque table, mais renferment également davantage de données pertinentes et intéressantes.

appercu d'un fichier DocumXML

```
library(xml2)

AviviaDBData$CommEW |> head(1) |> pull(DocumXML) |> as_xml_document() |> print()
```

```
{xml_document}
<xmlDunin version="1.00">
[1] <obj clas="Comm">\n  <prop nom="$aNote" type="boolean"><![CDATA[false]]>< ...
```

3.3 Transformation et enrichissement des données

3.3.1 structure de fichier xml

exemple de la structure des fichiers xmls (ils ont ensemble fondemmentalement la meme structure)

```
AviviaDBData$CommEW |> head(1) |> pull(DocumXML) |> as_xml_document() |> xml_structure()
```

```
<xmlDunin [version]>
  <obj [clas]>
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
      <carac [nom, type]>
        {cdata}
    <prop [nom, type]>
      {cdata}
      <carac [nom, type]>
        {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
```

```
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  <entr [no, type]>
    <obj [clas]>
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
```

```
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  <entr [no, type]>
    <obj [clas]>
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
  <entr [no, type]>
    <obj [clas]>
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
  <entr [no, type]>
    <obj [clas]>
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
```

```
        {cdata}
  <entr [no, type]>
    <obj [clas]>
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
      <prop [nom, type]>
        {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
  <carac [nom, type]>
    {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
  <notif [clas, ident, type]>
    {cdata}
<prop [nom, type]>
  {cdata}
```

```
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
```

```
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
<prop [nom, type]>
  {cdata}
```

```
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
    <prop [nom, type]>
      {cdata}
```

3.3.2 extractions des information et tables potentielles des différents fichiers xmls

3.3.2.1 fichiers DocumXML et EveXML de la table CommEW (commandes)

```
load("C:/Users/datch/Desktop/aviviaMandatDoc/data/commandes.RData")
#
commandes |> names()
```

```
 [1] "CabBase"            "CabBFiller"         "CabBFillerAuto"
 [4] "CabTable"           "CabTall"            "CabTFillerAuto"
 [7] "CabVanity"          "CabWall"            "CabWFiller"
[10] "CabWFillerAuto"     "CadreConfig"        "Comm"
[13] "Commentaire"        "Comptoir"           "Cooking"
[16] "DessousBT"          "DessousHotte"       "DessousWall"
[19] "Dessus"             "Disclaimer"         "Dishwashers"
[22] "DoorAPart"          "DoorConfig"         "DoorConfigAluVitre"
[25] "DoorConfigVitre"    "DoorLayout"         "Dos"
```

```
[28] "Etape"                "FillerTrav"           "FourConfig"
[31] "Ingenierie"           "Kit"                  "LayoutComptoir"
[34] "Lumiere"              "Molding"              "MoldingAssemble"
[37] "MoulLum"              "MoulLumGroupeAchat"   "MoulPied"
[40] "MoulPiedGroupeAchat" "MoulTete"             "MoulTeteGroupeAchat"
[43] "NicheVinConfig"       "NotifUser"            "PanBoxPlein"
[46] "PanBoxRect"           "PanDeco"              "PanelAPart"
[49] "PanelAuto"            "PanelCompos"          "PanelCote"
[52] "PanelDos"             "PanelFillerAuto"      "PanelWall"
[55] "PanelWallSous"        "PullOutConfig"        "Rangement2"
[58] "Refrigeration"        "RollOutConfig"        "SommaireKit"
[61] "TabletteConfig"       "TiroirConfig"         "TiroirConfigPIPN"
[64] "Various"              "Ventilation"          "WoodTop"
[67] "ZoneFaceApp"          "ZoneFaceAvant"        "ZoneFaceCote"
[70] "ZoneFaceDessous"      "ZoneFaceDessus"       "ZoneFaceDos"
[73] "ZoneFaceInt"          "ZoneSection"
```

3.3.2.2 fichiers DocumXML et EveXML de la table ContratEW (contrats)

```
load("C:/Users/datch/Desktop/aviviaMandatDoc/data/contrats.RData")
 #
 contrats |> names()
```

```
 [1] "CabBase"              "CabBFiller"           "CabBFillerAuto"
 [4] "CabTable"             "CabTall"              "CabTFiller"
 [7] "CabTFillerAuto"       "CabVanity"            "CabWall"
[10] "CabWFiller"           "CabWFillerAuto"       "CadreConfig"
[13] "Commentaire"          "ComposCustom"         "Comptoir"
[16] "Contrat"              "Cooking"              "DessousBT"
[19] "DessousHotte"         "DessousWall"          "Dessus"
[22] "Disclaimer"           "Dishwashers"          "DivisionConfig"
[25] "DoorAPart"            "DoorConfig"           "DoorConfigAluVitre"
[28] "DoorConfigVitre"      "DoorLayout"           "Dos"
[31] "Etape"                "FillerTrav"           "FillerTravSimple"
[34] "FourConfig"           "Ingenierie"           "Kit"
[37] "LayoutComptoir"       "Lumiere"              "MoldingAssemble"
[40] "MoulLumGroupeAchat"   "MoulPied"             "MoulPiedGroupeAchat"
[43] "MoulTete"             "MoulTeteGroupeAchat"  "NicheVinConfig"
[46] "NotifUser"            "PanBoxPlein"          "PanBoxRect"
[49] "PanBoxTrav"           "PanDeco"              "PanelAPart"
```

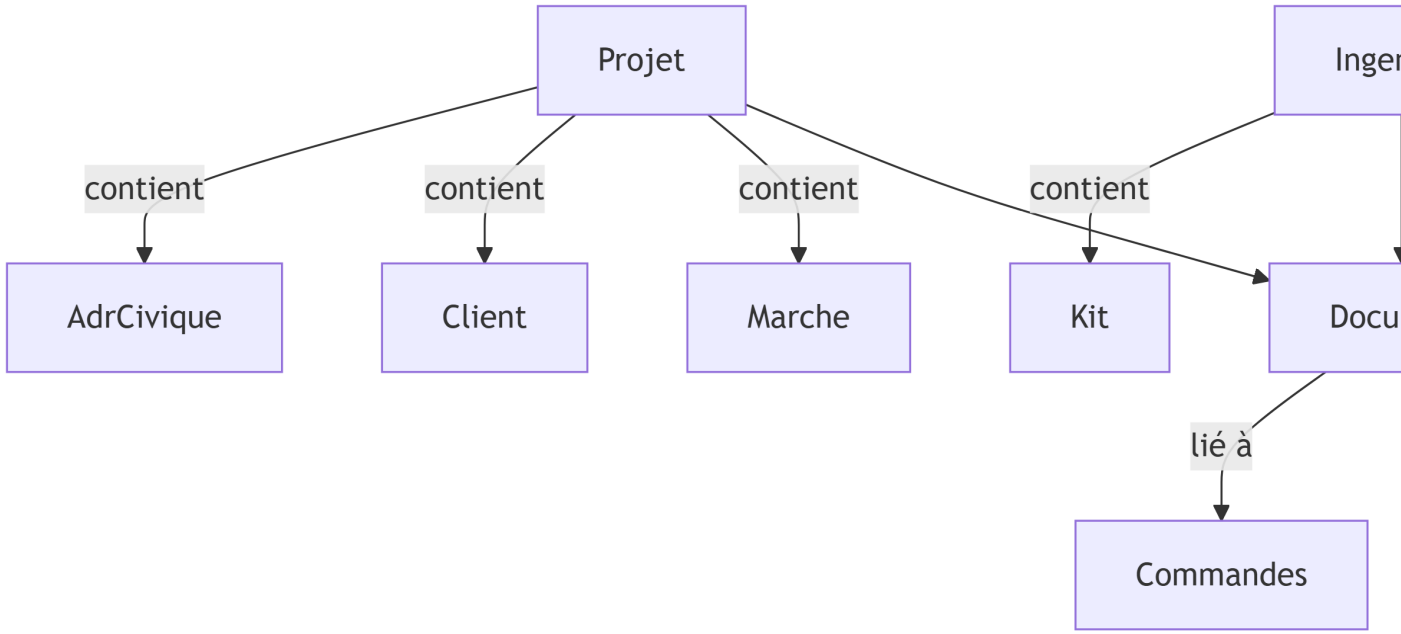
```
[52] "PanelAuto"           "PanelCompos"         "PanelCote"
[55] "PanelDos"            "PanelFillerAuto"     "PanelWall"
[58] "PanelWallSous"       "PullOutConfig"       "Rangement2"
[61] "Refrigeration"       "RollOutConfig"       "SommaireKit"
[64] "TabletteConfig"      "TiroirConfig"        "TiroirConfigPIPN"
[67] "Various"             "WasherDryers"        "WoodTop"
[70] "ZoneFaceApp"         "ZoneFaceAvant"       "ZoneFaceCote"
[73] "ZoneFaceDessous"     "ZoneFaceDessus"      "ZoneFaceDos"
[76] "ZoneFaceInt"         "ZoneSection"
```

3.3.2.3 fichiers ProjetXML de la table ProjetEw (projets)

```
load("C:/Users/datch/Desktop/aviviaMandatDoc/data/projets.RData")
#
projets |> names()
```

```
[1] "AdrCivique"    "AdrCourriel"   "Client"        "Commentaire"
[5] "CoordonneesJS" "Marche"        "Projet"        "Telephone"
```

3.3.3 diagrammmes de sélection des entités pour l'analyse



4 Analyses descriptives

4.1 Introduction

Cette section présente les statistiques descriptives des différentes entités analysées dans le projet: projet , ingenerie , documents.

4.2 Méthodologie

Statitstique descriptive univariée , bivariée et multivariée .

4.3 Structure de cette section

Cette section est divisée en plusieurs sous-sections, chacune se concentsrant sur une entité spécifique:

1. **documents (commandes , contrats)** : Analyse des commandes et contrats, volumes, tendances temporelles, etc.
2. **Projet(Clients, AdresseCivque, Marche)** : Analyse des clients, segmentation, comportements, etc.
3. **Ingenerie(Kit)** : Analyse des kits, compositions, popularité, etc.
4. **Synthèse**: Vue d'ensemble et synthèse des statistiques descriptives.

5 Analyse des documents (commandes)

Ce chapitre présente une analyse descriptive des commandes.

5.1 Vue d'ensemble des commandes et préparations des données

5.1.1 Résumé de la table

```
library(dplyr)
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
library(tidyr)
library(purrr)
load("C:/Users/datch/Desktop/aviviaMandatDoc/data/commandes.RData")

TableCommande <- commandes$Comm

TableCommande <- TableCommande |> select_if(~!all(is.na(.))) |> # supprimer les colonnes tota
              rename_with(~ gsub("[^A-Za-z0-9]", "", .))  |> #supprime les caractères spéciau
             rename_with(~tolower(.)) #convertir  les colonnes en format minuscules



TableCommande |> glimpse () # afficher le résumé de la table
```

```
Rows: 714
Columns: 191
$ anote                  <chr> "false", "", "false", "false", "false", "~
$ documsource            <chr> "false", "", "false", "false", "false", "~
$ estcq                  <chr> "false", "", "false", "false", "false", "~
$ estcqsup               <chr> "false", "", "false", "false", "false", "~
$ estimport              <chr> "false", "", "false", "false", "false", "~
$ estreprise             <chr> "false", "", "false", "false", "false", "~
$ acompte1               <chr> "305.7440%", "PSD40%", NA, NA, NA, "3322.~
$ acompte2               <chr> "420.3955%", "PSD55%", NA, NA, NA, "4568.~
$ dateacompte1emis       <chr> "2023/08/30", NA, NA, NA, NA, NA, NA, NA,~
$ fgfexclu               <chr> "false", "", "false", "false", "false", "~
$ installexclu           <chr> "false", "", "false", "false", "false", "~
$ matexclu               <chr> "false", "", "false", "false", "false", "~
$ moexclu                <chr> "false", "", "false", "false", "false", "~
$ pourcacompte1          <chr> "0", "0", "0", "0", "0", "0", "0", "0", "~
$ pourcacompte2          <chr> "0", "0", "0", "0", "0", "0", "0", "0", "~
$ solde                  <chr> "764.35", "PSD", NA, NA, NA, "8307.06", "~
$ transportexclu         <chr> "false", "", "false", "false", "false", "~
$ appnotifobligs         <chr> "false", "", "false", "false", "false", "~
$ appnotifs              <chr> "false", "", "false", "false", "false", "~
$ codeclient2            <chr> "Avivia", NA, "Avivia", "Avivia", "Avivia~
$ datevalidite           <chr> "2023/11/28", NA, NA, NA, NA, "2024/03/04~
$ delaiprevexped         <chr> "2", "2", "2", "2", "2", "2", "2", "2", "~
$ disclaimers            <chr> "", "", "", "", NA, "", "", "falsePriseMe~
$ emplacementcre         <chr> "Interne", "Interne", "Externe", "Externe~
$ enerreur               <chr> "false", "", "false", "false", "false", "~
$ enpsd                  <chr> "false", "", "false", "false", "false", "~
$ etapes                 <chr> "dunintech2023/08/30ManufacturierSaisie20~
$ identcour              <chr> "0", "0", "0", "0", "0", "0", "0", "0", "~
$ nbjourvalide           <chr> "90", "90", "90", "90", "90", "90", "90",~
$ nomclient              <chr> "Avivia", NA, "Avivia", "Avivia", "Avivia~
$ projet                 <chr> "P-000175", "P-000181", "P-000219", "P-00~
$ promotions             <chr> "", "", "", "", "", "", "", NA, "", "", "~
$ qteitems               <chr> "2", "4", "0", "0", "0", "10", "2", "26",~
$ qteitemscab            <chr> "1", "1", "0", "0", "0", "4", "1", "12", ~
$ roleproprietaire       <chr> "Employe", "Employe", "Client", "Client",~
$ sommaireskits          <chr> "k011Custom00346.43Contour000Plancher000P~
$ sorte                  <chr> "Comm", "Comm", "Comm", "Comm", "Comm", "~
$ statutcjobsite         <chr> "false", "", "false", "false", "false", "~
$ statutcred             <chr> "false", "", "false", "false", "false", "~
$ type                   <chr> "Regulier", "Regulier", "Regulier", "Regu~
$ typeprocess            <chr> "interne", "interne", "externe", "externe~
```

```
$ validcred                <chr> "true", "true", "true", "true", "true", "~
$ ligne                    <chr> "Custom", "Custom", "Stock", "Eco", NA, "~
$ avertlivraison           <chr> "Cout_Livraison_Peut_Modif", NA, NA, NA, ~
$ cleenmain                <chr> "false", "", "false", "false", "false", "~
$ codepost                 <chr> "G0A1L0", NA, "G0A1L0", "G0A1L0", "G0A1L0~
$ datedebvalidite          <chr> "2023-08-31true", NA, NA, NA, NA, "2023-1~
$ installation             <chr> "false", "", "false", "false", "false", "~
$ prisemesure              <chr> "falsepriseDeMesure", "nullpriseDeMesure"~
$ typeclient               <chr> "particulier", "particulier", "particulie~
$ assemblage               <chr> "assemble", "assemble", "nonAssemble", NA~
$ edgecolle                <chr> "pur", "pur", "pur", "pur", "pur", "pur",~
$ gamme                    <chr> "gamme_3", "gamme_2", "gamme_1", "gamme_3~
$ ciclient                 <chr> "0", NA, "0", "0", "0", "0", NA, NA, NA, ~
$ cleitem                  <chr> "24", "35", "7", "7", "1", "166", "25", "~
$ courrielcharge           <chr> "malavoie@dunin.com", NA, NA, NA, NA, NA,~
$ createurcourriel         <chr> "malavoie@dunin.com", "glebel@avivia.ca",~
$ createurnom              <chr> "Techdunin", "Ghislain", "client", "clien~
$ datecreation             <chr> "2023/08/30", "2023/08/31", "2023/09/14",~
$ division                 <chr> "Division_1", "Division_1", "Division_1",~
$ documcree                <chr> "1", "1", "1", "1", "1", "1", "1", "1", "~
$ memecoordsjscf           <chr> "false", "", "false", "false", "false", "~
$ sauvegarde               <chr> "true", "true", "false", "false", "false"~
$ statutjobsite            <chr> "false", "", "false", "false", "false", "~
$ code                     <chr> "C-000089", "C-000093", "C-000115", "C-00~
$ codetaxe                 <chr> "TPS", "N/A", "TPS", "TPS", "TPS", "TPS",~
$ codetaxe2                <chr> "TVQ", "N/A", "TVQ", "TVQ", "TVQ", "TVQ",~
$ fuelextra                <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
$ palettenet               <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
$ pourcesc                 <chr> "18", "0", "18", "18", "18", "18", "0", "~
$ pourcescsup              <chr> "0", "0", NA, NA, NA, NA, "0", NA, "0", N~
$ soustotal                <chr> "346.43", "330.66", "0", "0", "0", "1935.~
$ tauxtaxe                 <chr> "5", "0", "5", "5", "5", "5", "0", "5", "~
$ tauxtaxe2                <chr> "9.975", "0", "9.975", "9.975", "9.975", ~
$ total                    <chr> "764.35", "PSD", NA, NA, NA, "8307.06", "~
$ totalht                  <chr> "664.8", "PSD", NA, NA, NA, "7225.1", "34~
$ totaltx1                 <chr> "33.24", NA, NA, NA, NA, "361.26", NA, "6~
$ totaltx2                 <chr> "66.31", NA, NA, NA, NA, "720.7", NA, "12~
$ transpcumul              <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
$ transpextra              <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
$ transpjobsite            <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
$ transpnet                <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
$ aveclivraison            <chr> "true", "1", "1", "1", "1", "1", "1", "1"~
$ cumulservice             <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
```

```
$ cumulvolume                <chr> "4.36", "4.36", NA, NA, NA, "31.37", "4.3~
$ pourcescautre              <chr> "0", "0", "0", "0", "0", "0", "0", "0", "~
$ pourcesccab                <chr> "0", "0", "0", "0", "0", "0", "0", "0", "~
$ pourcesccomptoir           <chr> "0", "0", "0", "0", "0", "0", "0", "0", "~
$ pourcescmoulure            <chr> "0", "0", "0", "0", "0", "0", "0", "0", "~
$ pourcescpanel              <chr> "0", "0", "0", "0", "0", "0", "0", "0", "~
$ soustotalautre             <chr> "0", "", "0", "0", "0", "0", "6.03", "408~
$ soustotalcab               <chr> "341.02", "297.37", "0", "0", "0", "1519.~
$ soustotalcomptoir          <chr> "0", "PSD", "0", "0", "0", NA, "0", "2701~
$ soustotalmoulure           <chr> "5.41", "11.86", "0", "0", "0", "17.29", ~
$ soustotalpanel             <chr> "0", "21.43", "0", "0", "0", "6.87", "0",~
$ totalcomptoirbois          <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
$ totalcomptoirgranite       <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
$ totalcomptoirquartz        <chr> "0", "PSD", "0", "0", "0", NA, "0", "2701~
$ totalcomptoirstratifie     <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
$ totalescautre              <chr> "0", "", "0", "0", "0", "0", NA, NA, "0",~
$ totalesccomptoir           <chr> "0", "", "0", "0", "0", NA, "0", NA, "0",~
$ totalescpanel              <chr> "0", NA, "0", "0", "0", NA, "0", NA, "0",~
$ totallivraison             <chr> "318.37", "PSD", NA, NA, NA, "289.55", "P~
$ unit                       <chr> "0", "", "0", "0", "0", "0", "0", "0", "0~
$ commandes                  <chr> "C-000089", "C-000093", "C-000115", "C-00~
$ soustotal2                 <chr> NA, "PSD", NA, NA, NA, "6935.55", "348.28~
$ ecartprix                  <chr> NA, "", NA, NA, NA, "false", "false", "fa~
$ fraispsd                   <chr> NA, "PSD_LIV", NA, NA, NA, NA, "PSD_LIV",~
$ montantecartprix           <chr> NA, "", NA, NA, NA, "0", "0", "0", NA, "0~
$ soustotal2clientfinal      <chr> NA, "PSD", NA, NA, NA, NA, NA, NA, NA, NA~
$ soustotalcomptoirdealer    <chr> NA, "PSD", NA, NA, NA, NA, NA, NA, NA, NA~
$ soustotalservice           <chr> NA, "PSD", NA, NA, NA, NA, NA, NA, NA, NA~
$ soustotalserviceesc        <chr> NA, "PSD", NA, NA, NA, NA, NA, NA, NA, NA~
$ soustotaldealer            <chr> NA, "330.66", NA, NA, NA, NA, NA, NA, NA,~
$ totalescservice            <chr> NA, "", NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalhtcf                  <chr> NA, "PSD", NA, NA, NA, NA, NA, NA, NA, NA~
$ totalcf                    <chr> NA, "PSD", NA, NA, NA, NA, NA, NA, NA, NA~
$ soustotal2eff              <chr> NA, "330.66", NA, NA, NA, "2327.03", "348~
$ soustotalcabeff            <chr> NA, "297.37", NA, NA, NA, "1519.91", "342~
$ soustotaleff               <chr> NA, "330.66", NA, NA, NA, "1935.55", "348~
$ soustotalmoulureeff        <chr> NA, "11.86", NA, NA, NA, "17.29", "0", "0~
$ soustotalpaneleff          <chr> NA, "21.43", NA, NA, NA, "6.87", "0", "64~
$ soustotaldealereff         <chr> NA, "330.66", NA, NA, NA, NA, NA, NA, NA,~
$ totalesccab                <chr> NA, NA, "0", "0", "0", NA, NA, NA, "0", N~
$ totalescmoulure            <chr> NA, NA, "0", "0", "0", NA, "0", "0", "0",~
$ gammedesc                  <chr> NA, NA, NA, NA, NA, "gamme_desc_2true", "~
$ escsup                     <chr> NA, NA, NA, NA, NA, "0", "0", "0", NA, "0~
```

```
$ estbloqprix                  <chr> NA, NA, NA, NA, NA, "false", "false", "fa~
$ soustotalautreeff            <chr> NA, NA, NA, NA, NA, "0", "6.03", "408.14"~
$ soustotalcomptoireff         <chr> NA, NA, NA, NA, NA, "733.42", "0", "2701.~
$ totalcomptoirboiseff         <chr> NA, NA, NA, NA, NA, "0", "0", "0", NA, "0~
$ totalcomptoirgraniteeff      <chr> NA, NA, NA, NA, NA, "0", "0", "0", NA, "0~
$ totalcomptoirquartzeff       <chr> NA, NA, NA, NA, NA, "733.42", "0", "2701.~
$ totalcomptoirstratifieeff    <chr> NA, NA, NA, NA, NA, "0", "0", "0", NA, "0~
$ totalescautreeff             <chr> NA, NA, NA, NA, NA, "0", NA, NA, NA, "0",~
$ signaturevalidation          <chr> NA, NA, NA, NA, NA, NA, "false", "false",~
$ memecoordsjs                 <chr> NA, NA, NA, NA, NA, NA, "false", "false",~
$ totalesccomptoireff          <chr> NA, NA, NA, NA, NA, NA, "0", NA, NA, NA, ~
$ totalescmoulureeff           <chr> NA, NA, NA, NA, NA, NA, "0", "0", NA, NA,~
$ totalescpaneleff             <chr> NA, NA, NA, NA, NA, NA, "0", NA, NA, "0",~
$ installateur                 <chr> NA, NA, NA, NA, NA, NA, NA, "particuliers~
$ lumieres                     <chr> NA, NA, NA, NA, NA, NA, NA, "PUCK79mm_WH3~
$ ref1                         <chr> NA, NA, NA, NA, NA, NA, NA, "cuisine 2tru~
$ codepostfinal                <chr> NA, NA, NA, NA, NA, NA, NA, "G0A1L0", NA,~
$ cumulinstall                 <chr> NA, NA, NA, NA, NA, NA, NA, "1199.77", NA~
$ fraispdm                     <chr> NA, NA, NA, NA, NA, NA, NA, "120", NA, NA~
$ totalinstall                 <chr> NA, NA, NA, NA, NA, NA, NA, "1499.77", NA~
$ totalesc                     <chr> NA, NA, NA, NA, NA, NA, NA, NA, "0", NA, ~
$ totalesccabeff               <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, "~
$ totalesceff                  <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, "~
$ contratsource                <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ datereqexped                 <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totaltx1cf                   <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totaltx2cf                   <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ soustotalcomptoirdealereff   <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ dateacompte1recep            <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ courrielclient               <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ commentaires                 <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ nomcharge                    <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ telephonecharge              <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ soustotalautredealer         <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ soustotalautredealereff      <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalesccomptoirboiseff      <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalesccomptoirgraniteeff   <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalesccomptoirquartzeff    <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalesccomptoirstratifieeff <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ condition01                  <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ condition02                  <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ condition03                  <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ condition04                  <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
```

```
$ condition05                  <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ condition06                  <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ cumulcem                     <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ fraissav                     <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ dateprisemesure              <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ emailplanifenvoi             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ dateprevexped                <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalcem                     <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ specsportes                  <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ nomclientfin                 <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ prenomclient                 <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ telephoneclient              <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ dateappclient                <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ dateacompte2emis             <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ dateacompte2recep            <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalescompte                <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalesccomptoirbois         <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalesccomptoirgranite      <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalesccomptoirquartz       <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ totalesccomptoirstratifie    <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ datelivraisoninstall         <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
```

5.1.2 selections des variables (colonnes pertinentes)

```
variables_utiles <- c(
  "nomclient", "codeclient2", "typeclient", "ciclient",
  "datecreation", "datevalidite", "datedebvalidite", "datereqexped", "dateprevexped", "datel
  "acompte1", "acompte2", "solde", "total", "totalht",
  "soustotal", "soustotalcab", "soustotalcomptoir", "soustotalservice", "totallivraison",
  "qteitems", "qteitemscab", "gamme", "gammedesc",
  "statutcred", "statutcjobsite", "validcred",
  "aveclivraison", "cleenmain", "prisemesure", "installation",
  "codepost", "emplacementcre", "division",
  "commentaires", "courrielclient", "nomcharge", "telephoneclient"
)


TableCommande <- TableCommande[,variables_utiles]

TableCommande |> glimpse()
```

```
Rows: 714
Columns: 39
$ nomclient            <chr> "Avivia", NA, "Avivia", "Avivia", "Avivia", "Aviv~
$ codeclient2          <chr> "Avivia", NA, "Avivia", "Avivia", "Avivia", "Aviv~
$ typeclient           <chr> "particulier", "particulier", "particulier", "par~
$ ciclient             <chr> "0", NA, "0", "0", "0", "0", NA, NA, NA, "0", NA,~
$ datecreation         <chr> "2023/08/30", "2023/08/31", "2023/09/14", "2023/0~
$ datevalidite         <chr> "2023/11/28", NA, NA, NA, NA, "2024/03/04", NA, N~
$ datedebvalidite      <chr> "2023-08-31true", NA, NA, NA, NA, "2023-12-06true~
$ datereqexped         <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, "~
$ dateprevexped        <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ datelivraisoninstall <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ acompte1             <chr> "305.7440%", "PSD40%", NA, NA, NA, "3322.8240%", ~
$ acompte2             <chr> "420.3955%", "PSD55%", NA, NA, NA, "4568.8855%", ~
$ solde                <chr> "764.35", "PSD", NA, NA, NA, "8307.06", "348.28",~
$ total                <chr> "764.35", "PSD", NA, NA, NA, "8307.06", "348.28",~
$ totalht              <chr> "664.8", "PSD", NA, NA, NA, "7225.1", "348.28", "~
$ soustotal            <chr> "346.43", "330.66", "0", "0", "0", "1935.55", "34~
$ soustotalcab         <chr> "341.02", "297.37", "0", "0", "0", "1519.91", "34~
$ soustotalcomptoir    <chr> "0", "PSD", "0", "0", "0", NA, "0", "2701.76", "0~
$ soustotalservice     <chr> NA, "PSD", NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ totallivraison       <chr> "318.37", "PSD", NA, NA, NA, "289.55", "PSD", "28~
$ totalinstall         <chr> NA, NA, NA, NA, NA, NA, NA, "1499.77", NA, NA, NA~
$ qteitems             <chr> "2", "4", "0", "0", "0", "10", "2", "26", "0", "8~
$ qteitemscab          <chr> "1", "1", "0", "0", "0", "4", "1", "12", "0", "3"~
$ gamme                <chr> "gamme_3", "gamme_2", "gamme_1", "gamme_3", NA, "~
$ gammedesc            <chr> NA, NA, NA, NA, NA, "gamme_desc_2true", "gamme_de~
$ statutcred           <chr> "false", "", "false", "false", "false", "false", ~
$ statutcjobsite       <chr> "false", "", "false", "false", "false", "false", ~
$ validcred            <chr> "true", "true", "true", "true", "true", "true", "~
$ aveclivraison        <chr> "true", "1", "1", "1", "1", "1", "1", "1", "true"~
$ cleenmain            <chr> "false", "", "false", "false", "false", "false", ~
$ prisemesure          <chr> "falsepriseDeMesure", "nullpriseDeMesure", "false~
$ installation         <chr> "false", "", "false", "false", "false", "false", ~
$ codepost             <chr> "G0A1L0", NA, "G0A1L0", "G0A1L0", "G0A1L0", "G0A1~
$ emplacementcre       <chr> "Interne", "Interne", "Externe", "Externe", "Exte~
$ division             <chr> "Division_1", "Division_1", "Division_1", "Divisi~
$ commentaires         <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ courrielclient       <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ nomcharge            <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ telephoneclient      <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
```

5.1.3 Conversion des colones dans des types de variables appropriés

```
library(dplyr)
library(stringr)
library(lubridate)
```

```
Attaching package: 'lubridate'

The following objects are masked from 'package:base':

    date, intersect, setdiff, union
```

```
TableCommande_cleaned <- TableCommande |>
  mutate(
    # Conversion des dates
    datecreation = ymd(datecreation),
    datevalidite = ymd(datevalidite),
    datedebvalidite = ymd(str_remove(datedebvalidite, "true")),
    datereqexped = ymd(datereqexped),
    dateprevexped = ymd(dateprevexped),
    datelivraisoninstall = ymd(datelivraisoninstall),

    # Garder les acomptes en double mais nettoyer leur format
    acompte1 = str_replace(acompte1, "%", ""), # Enlever le % s'il existe
    acompte1 = as.numeric(str_extract(acompte1, "[0-9.]+")), # Extraire la valeur numérique
    acompte2 = str_replace(acompte2, "%", ""), # Enlever le % s'il existe
    acompte2 = as.numeric(str_extract(acompte2, "[0-9.]+")), # Extraire la valeur numérique

    # Nettoyer la colonne prisemesure et convertir en booléen
    prisemesure = case_when(
      str_detect(prisemesure, "false") ~ FALSE,
      str_detect(prisemesure, "true") ~ TRUE,
      str_detect(prisemesure, "null") ~ NA,
      TRUE ~ NA
    ),

    # Conversion des autres colonnes numériques
    across(c(solde, total, totalht, soustotal, soustotalcab,
             soustotalcomptoir, totallivraison, totalinstall),
```

```
          ~ as.numeric(str_replace_all(., "[^0-9.-]", ""))),

    # Conversion des booléens
    across(c(statutcred, statutcjobsite, validcred, aveclivraison, cleenmain, installation),
           ~ case_when(
             . %in% c("true", "1") ~ TRUE,
             . %in% c("false", "", "0") ~ FALSE,
             TRUE ~ NA
           )),

    # Conversion des autres types
    across(c(qteitems, qteitemscab), as.integer),
    across(c(nomclient, codeclient2, typeclient, gamme, gammedesc, emplacementcre, division)
  )
```

```
Warning: There were 3 warnings in `mutate()`.
The first warning was:
i In argument: `datevalidite = ymd(datevalidite)`.
Caused by warning:
!  1 failed to parse.
i Run `dplyr::last_dplyr_warnings()` to see the 2 remaining warnings.
```

```
#

TableCommande_cleaned |> glimpse()
```

```
Rows: 714
Columns: 39
$ nomclient            <fct> Avivia, NA, Avivia, Avivia, Avivia, Avivia, NA, A~
$ codeclient2          <fct> Avivia, NA, Avivia, Avivia, Avivia, Avivia, NA, A~
$ typeclient           <fct> particulier, particulier, particulier, particulie~
$ ciclient             <chr> "0", NA, "0", "0", "0", "0", NA, NA, NA, "0", NA,~
$ datecreation         <date> 2023-08-30, 2023-08-31, 2023-09-14, 2023-09-14, ~
$ datevalidite         <date> 2023-11-28, NA, NA, NA, NA, 2024-03-04, NA, NA, ~
$ datedebvalidite      <date> 2023-08-31, NA, NA, NA, NA, 2023-12-06, NA, NA, ~
$ datereqexped         <date> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ dateprevexped        <date> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ datelivraisoninstall <date> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ acompte1             <dbl> 305.744, 40.000, NA, NA, NA, 3322.824, 139.314, 5~
$ acompte2             <dbl> 420.3955, 55.0000, NA, NA, NA, 4568.8855, 191.555~
$ solde                <dbl> 764.35, NA, NA, NA, NA, 8307.06, 348.28, 13918.37~
```

```
$ total                <dbl> 764.35, NA, NA, NA, NA, 8307.06, 348.28, 13918.37~
$ totalht              <dbl> 664.80, NA, NA, NA, NA, 7225.10, 348.28, 12105.56~
$ soustotal            <dbl> 346.43, 330.66, 0.00, 0.00, 0.00, 1935.55, 348.28~
$ soustotalcab         <dbl> 341.02, 297.37, 0.00, 0.00, 0.00, 1519.91, 342.25~
$ soustotalcomptoir    <dbl> 0.00, NA, 0.00, 0.00, 0.00, NA, 0.00, 2701.76, 0.~
$ soustotalservice     <chr> NA, "PSD", NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ totallivraison       <dbl> 318.37, NA, NA, NA, NA, 289.55, NA, 289.55, NA, 3~
$ totalinstall         <dbl> NA, NA, NA, NA, NA, NA, NA, 1499.77, NA, NA, NA, ~
$ qteitems             <int> 2, 4, 0, 0, 0, 10, 2, 26, 0, 8, 0, 0, 6, 2, 0, 2,~
$ qteitemscab          <int> 1, 1, 0, 0, 0, 4, 1, 12, 0, 3, 0, 0, 2, 1, 0, 1, ~
$ gamme                <fct> gamme_3, gamme_2, gamme_1, gamme_3, NA, gamme_2, ~
$ gammedesc            <fct> NA, NA, NA, NA, NA, gamme_desc_2true, gamme_desc_~
$ statutcred           <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ statutcjobsite       <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ validcred            <lgl> TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, T~
$ aveclivraison        <lgl> TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, T~
$ cleenmain            <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ prisemesure          <lgl> FALSE, NA, FALSE, FALSE, FALSE, FALSE, FALSE, TRU~
$ installation         <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ codepost             <chr> "G0A1L0", NA, "G0A1L0", "G0A1L0", "G0A1L0", "G0A1~
$ emplacementcre       <fct> Interne, Interne, Externe, Externe, Externe, Inte~
$ division             <fct> Division_1, Division_1, Division_1, Division_1, D~
$ commentaires         <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ courrielclient       <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ nomcharge            <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ telephoneclient      <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
```

5.1.4 Gestion des valeurs manquantes (NA)

```
library(naniar)
library(ggplot2)
```

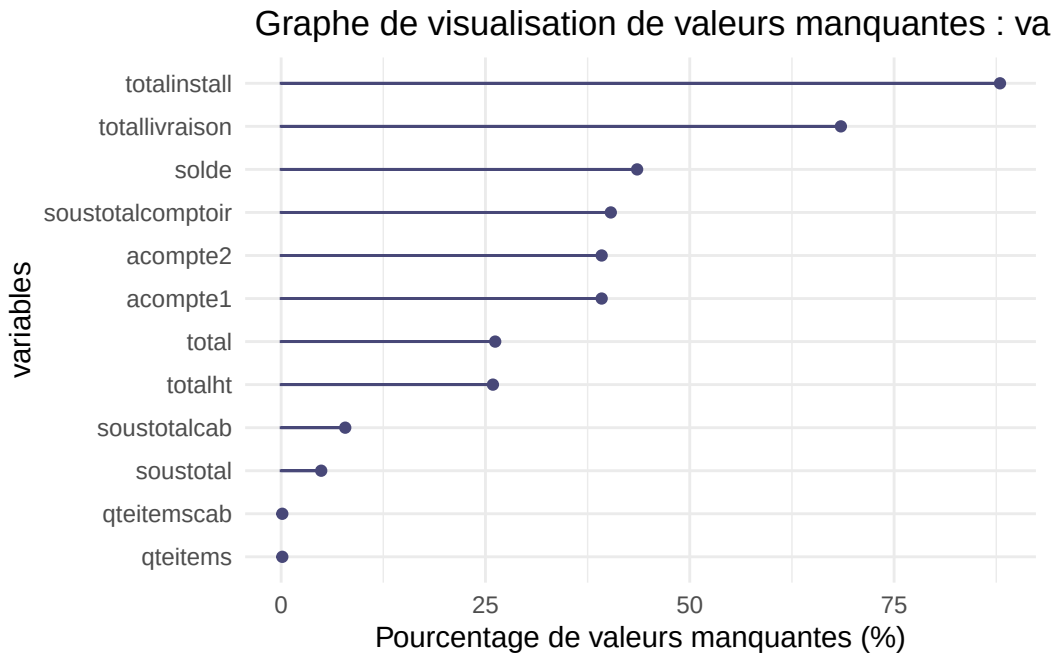
5.1.4.1 variables quantitatives

5.1.4.1.1 graphe

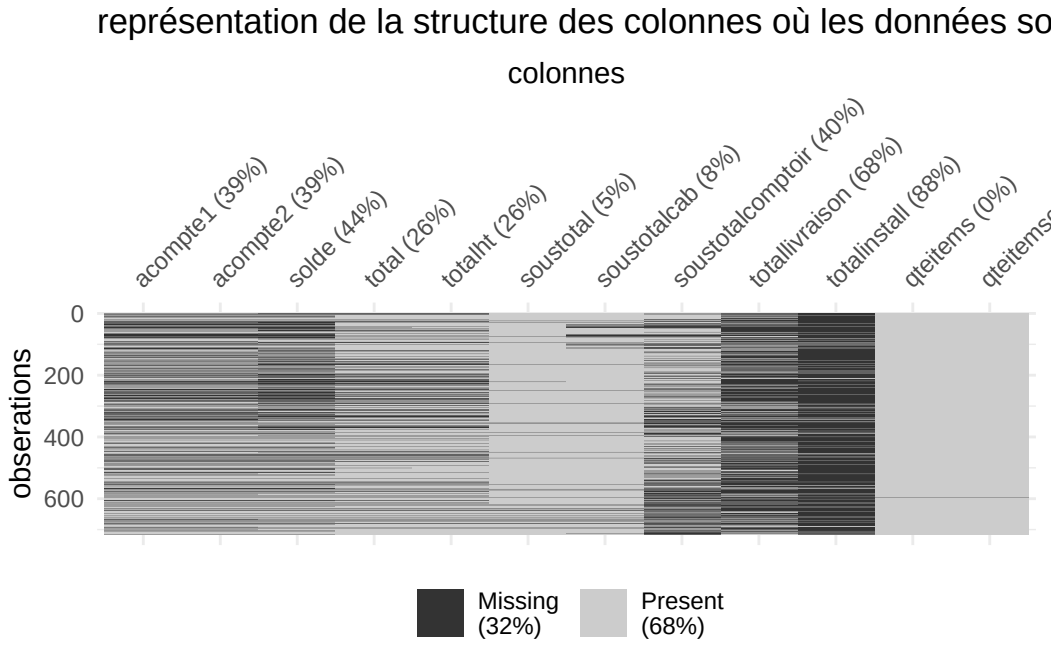
```
varsQuantitativeCommandes <- TableCommande_cleaned |> select(where(is.numeric))

# visualer le taux des valeurs manquantes par variable
```

```r
# graphe 1
gg_miss_var(varsQuantitativeCommandes,show_pct = TRUE) + labs(title = " Graphe de visualisati
```




```r
# graphe 2
naniar::vis_miss(varsQuantitativeCommandes) + labs(
  title = "représentation de la structure des colonnes où les données sont absentes.", y = "
)
```




```
# suppression des variable avec plus 50% de valeurs manquantes totalinstall , total livraison
varsQuantitativeCommandes <- varsQuantitativeCommandes |>
  mutate(totallivraison =NULL,totalinstall =NULL) |>
  arrange(qteitems) # ordonner de la plus petite à la plus grande quantite items pour la suit
```

5.1.4.1.2 Vérification MCAR (Missing Completely at Random) :

Rappel :

- MCAR (Missing Completely at Random) signifie que l'absence de données est complètement aléatoire et n'est liée à aucune autre caractéristique ou variable de l'ensemble de données. voir plus ...
- **Test de Litte** :
  1. Une p-valeur élevée (> 0.05) indique que les données sont probablement MCAR.
  2. Une p-valeur faible (< 0.05) suggère que les données sont MAR (Missing At Random) ou MNAR (Missing Not At Random)

```
LittLeTestResultat <- mcar_test(varsQuantitativeCommandes)
LittLeTestResultat |> print()
```

```
# A tibble: 1 x 4
  statistic    df p.value missing.patterns
      <dbl> <dbl>   <dbl>            <int>
1      581.   100       0               20
```

conclusion : l'absence de données est liée aux caractéristiques ou variables de l'ensemble de données

5.1.4.1.3 solution d'imputation de valeurs k-nearest neighbors (knn) :

Pour chaque observation avec des valeurs manquantes, identifier les k observations les plus proches parmi celles qui ont des valeurs complètes pour les variables concernées. Utiliser les valeurs des voisins (la moyenne (ou la médiane)) pour imputer la valeur manquante :

```
library(VIM)
```

```
Loading required package: colorspace
```

```
Loading required package: grid
```

```
VIM is ready to use.
```

```
Suggestions and bug-reports can be submitted at: https://github.com/statistikat/VIM/issues
```

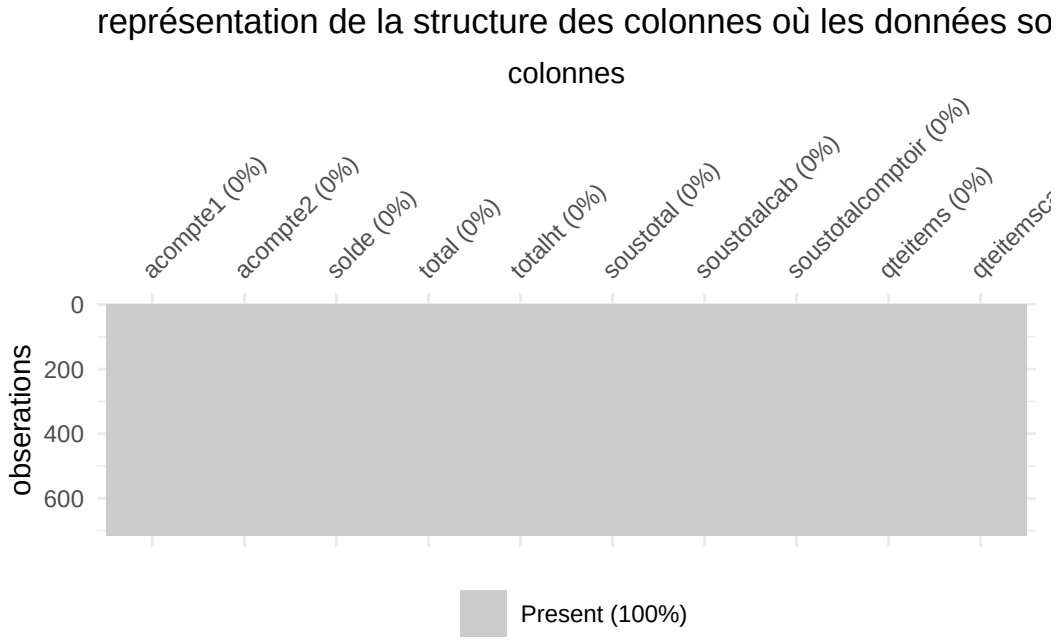
```
Attaching package: 'VIM'
```

```
The following object is masked from 'package:datasets':

    sleep
```

```
result <- kNN(varsQuantitativeCommandes,k=7)
varsQuantitativeCommandes <-   result[, 1:ncol(varsQuantitativeCommandes)]

# graphe 2
naniar::vis_miss(varsQuantitativeCommandes) + labs(
  title = "représentation de la structure des colonnes où les données sont absentes.", y = "
)
```




```
# resumé statistique des différentes colones
varsQuantitativeCommandes |> summary()
```

```
   acompte1           acompte2            solde              total        
 Min.   :     0.6   Min.   :     0.8   Min.   :    1.42   Min.   :    0.0  
 1st Qu.:    50.0   1st Qu.:    55.0   1st Qu.:  953.86   1st Qu.:  415.8  
 Median :   395.2   Median :   503.0   Median : 1836.16   Median : 1836.2  
 Mean   :  3763.7   Mean   :  4594.5   Mean   : 4821.90   Mean   : 4933.5  
 3rd Qu.:  2719.5   3rd Qu.:  3003.9   3rd Qu.: 7001.22   3rd Qu.: 7692.1  
 Max.   :625140.0   Max.   :623655.0   Max.   :35947.12   Max.   :35947.1  
   totalht         soustotal         soustotalcab      soustotalcomptoir
 Min.   :    0   Min.   :    0.0   Min.   :    0.0   Min.   :   0.0     
 1st Qu.:  180   1st Qu.:    0.0   1st Qu.:    0.0   1st Qu.:   0.0     
 Median : 1353   Median :  807.8   Median :  650.7   Median : 272.9     
 Mean   : 4056   Mean   : 3222.3   Mean   : 2847.6   Mean   : 790.8     
 3rd Qu.: 5814   3rd Qu.: 4763.9   3rd Qu.: 4330.3   3rd Qu.: 942.5     
 Max.   :31265   Max.   :25546.2   Max.   :21425.9   Max.   :8367.7     
   qteitems       qteitemscab     
 Min.   : 0.00   Min.   : 0.000  
 1st Qu.: 0.00   1st Qu.: 0.000  
 Median : 4.00   Median : 1.000  
 Mean   :11.26   Mean   : 5.662  
```

```
 3rd Qu.:18.00   3rd Qu.: 9.000
 Max.   :72.00   Max.   :37.000
```

5.2 Distribution des commandes par montant

[Analyse de la distribution des commandes par montant]

5.3 Évolution temporelle des commandes

[Analyse de l'évolution temporelle des commandes]

5.4 Commandes par canal

[Analyse de la répartition des commandes par canal]

5.5 Commandes par région

[Analyse de la répartition des commandes par région]

5.6 Taux de conversion

[Analyse des taux de conversion]

5.7 Synthèse de l'analyse des commandes

[Synthèse des principales observations et conclusions concernant les commandes]

Part III

Analyse avancée

6 Titre à remplir

7 À compléter

Contenu à rédiger ici.

Part IV

Application finale

8 Titre à remplir

9 À compléter

Contenu à rédiger ici.