rapport

Suibgui khaireddine

2024-05-11

Analyse des réservations de logements pour une région

Chargement des bibliothèques

```
library(dplyr)
library(tidyverse)
library(magrittr)
library(readr)
library(ggplot2)
library(randomForest)
library(caret)
library(lattice)
library(glmnet)
library(tinytex)
```

Importation de la base de données et affichage des variables

```
data <- read.csv("data_base.csv")
names(data)
```

```
##  [1] "id"                            "name"
##  [3] "host_id"                       "host_name"
##  [5] "neighbourhood_group"           "neighbourhood"
##  [7] "latitude"                      "longitude"
##  [9] "room_type"                     "price"
## [11] "minimum_nights"                "number_of_reviews"
## [13] "last_review"                   "reviews_per_month"
## [15] "calculated_host_listings_count" "availability_365"
```

affichage de lignes de la base de données

```
head(data)
```

```
##     id                                              name host_id   host_name
## 1 2539              Clean & quiet apt home by the park    2787        John
## 2 2595                        Skylit Midtown Castle        2845    Jennifer
## 3 3647            THE VILLAGE OF HARLEM....NEW YORK !      4632    Elisabeth
## 4 3831               Cozy Entire Floor of Brownstone      4869  LisaRoxanne
## 5 5022 Entire Apt: Spacious Studio/Loft by central park    7192        Laura
## 6 5099        Large Cozy 1 BR Apartment In Midtown East    7322        Chris
```

```
##   neighbourhood_group neighbourhood latitude longitude       room_type price
## 1            Brooklyn    Kensington 40.64749 -73.97237    Private room   149
## 2           Manhattan       Midtown 40.75362 -73.98377 Entire home/apt   225
## 3           Manhattan        Harlem 40.80902 -73.94190    Private room   150
## 4            Brooklyn  Clinton Hill 40.68514 -73.95976 Entire home/apt    89
## 5           Manhattan   East Harlem 40.79851 -73.94399 Entire home/apt    80
## 6           Manhattan   Murray Hill 40.74767 -73.97500 Entire home/apt   200
##   minimum_nights number_of_reviews last_review reviews_per_month
## 1              1                 9  2018-10-19              0.21
## 2              1                45  2019-05-21              0.38
## 3              3                 0                            NA
## 4              1               270  2019-07-05              4.64
## 5             10                 9  2018-11-19              0.10
## 6              3                74  2019-06-22              0.59
##   calculated_host_listings_count availability_365
## 1                              6              365
## 2                              2              355
## 3                              1              365
## 4                              1              194
## 5                              1                0
## 6                              1              129
```

Suppression des lignes avec des valeurs manquantes

```
data_clean <- na.omit(data)
head(data_clean)
```

```
##     id                                             name host_id   host_name
## 1 2539                 Clean & quiet apt home by the park    2787        John
## 2 2595                             Skylit Midtown Castle    2845    Jennifer
## 4 3831                   Cozy Entire Floor of Brownstone    4869 LisaRoxanne
## 5 5022 Entire Apt: Spacious Studio/Loft by central park    7192       Laura
## 6 5099       Large Cozy 1 BR Apartment In Midtown East    7322       Chris
## 7 5121                                   BlissArtsSpace!    7356       Garon
##   neighbourhood_group      neighbourhood latitude longitude       room_type
## 1            Brooklyn         Kensington 40.64749 -73.97237    Private room
## 2           Manhattan            Midtown 40.75362 -73.98377 Entire home/apt
## 4            Brooklyn       Clinton Hill 40.68514 -73.95976 Entire home/apt
## 5           Manhattan        East Harlem 40.79851 -73.94399 Entire home/apt
## 6           Manhattan        Murray Hill 40.74767 -73.97500 Entire home/apt
## 7            Brooklyn Bedford-Stuyvesant 40.68688 -73.95596    Private room
##   price minimum_nights number_of_reviews last_review reviews_per_month
## 1   149              1                 9  2018-10-19              0.21
## 2   225              1                45  2019-05-21              0.38
## 4    89              1               270  2019-07-05              4.64
## 5    80             10                 9  2018-11-19              0.10
## 6   200              3                74  2019-06-22              0.59
## 7    60             45                49  2017-10-05              0.40
##   calculated_host_listings_count availability_365
## 1                              6              365
## 2                              2              355
## 4                              1              194
## 5                              1                0
```

2

```
## 6                                 1           129
## 7                                 1             0
```

Suppression des doublons

```
data_clean <- distinct(data_clean)
head(data_clean)
```

```
##      id                                          name host_id   host_name
## 1 2539             Clean & quiet apt home by the park    2787        John
## 2 2595                          Skylit Midtown Castle    2845    Jennifer
## 3 3831                Cozy Entire Floor of Brownstone    4869  LisaRoxanne
## 4 5022 Entire Apt: Spacious Studio/Loft by central park  7192       Laura
## 5 5099        Large Cozy 1 BR Apartment In Midtown East   7322       Chris
## 6 5121                                BlissArtsSpace!    7356       Garon
##   neighbourhood_group       neighbourhood latitude longitude       room_type
## 1            Brooklyn          Kensington 40.64749 -73.97237    Private room
## 2           Manhattan             Midtown 40.75362 -73.98377 Entire home/apt
## 3            Brooklyn        Clinton Hill 40.68514 -73.95976 Entire home/apt
## 4           Manhattan         East Harlem 40.79851 -73.94399 Entire home/apt
## 5           Manhattan         Murray Hill 40.74767 -73.97500 Entire home/apt
## 6            Brooklyn Bedford-Stuyvesant 40.68688 -73.95596    Private room
##   price minimum_nights number_of_reviews last_review reviews_per_month
## 1   149              1                 9  2018-10-19              0.21
## 2   225              1                45  2019-05-21              0.38
## 3    89              1               270  2019-07-05              4.64
## 4    80             10                 9  2018-11-19              0.10
## 5   200              3                74  2019-06-22              0.59
## 6    60             45                49  2017-10-05              0.40
##   calculated_host_listings_count availability_365
## 1                              6              365
## 2                              2              355
## 3                              1              194
## 4                              1                0
## 5                              1              129
## 6                              1                0
```

Affichage des statistiques descriptives

```
summary(data_clean$price)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0    69.0   101.0   142.3   170.0 10000.0
```

Calcul de la moyenne et de l'écart type des reviews

```
summarized_data <- data %>%
  summarize(mean_reviews = mean(number_of_reviews),
          sd_reviews = sd(number_of_reviews),
          median_reviews = median(number_of_reviews))
head(summarized_data)
```

```
##   mean_reviews sd_reviews median_reviews
## 1     23.27447   44.55058              5
```
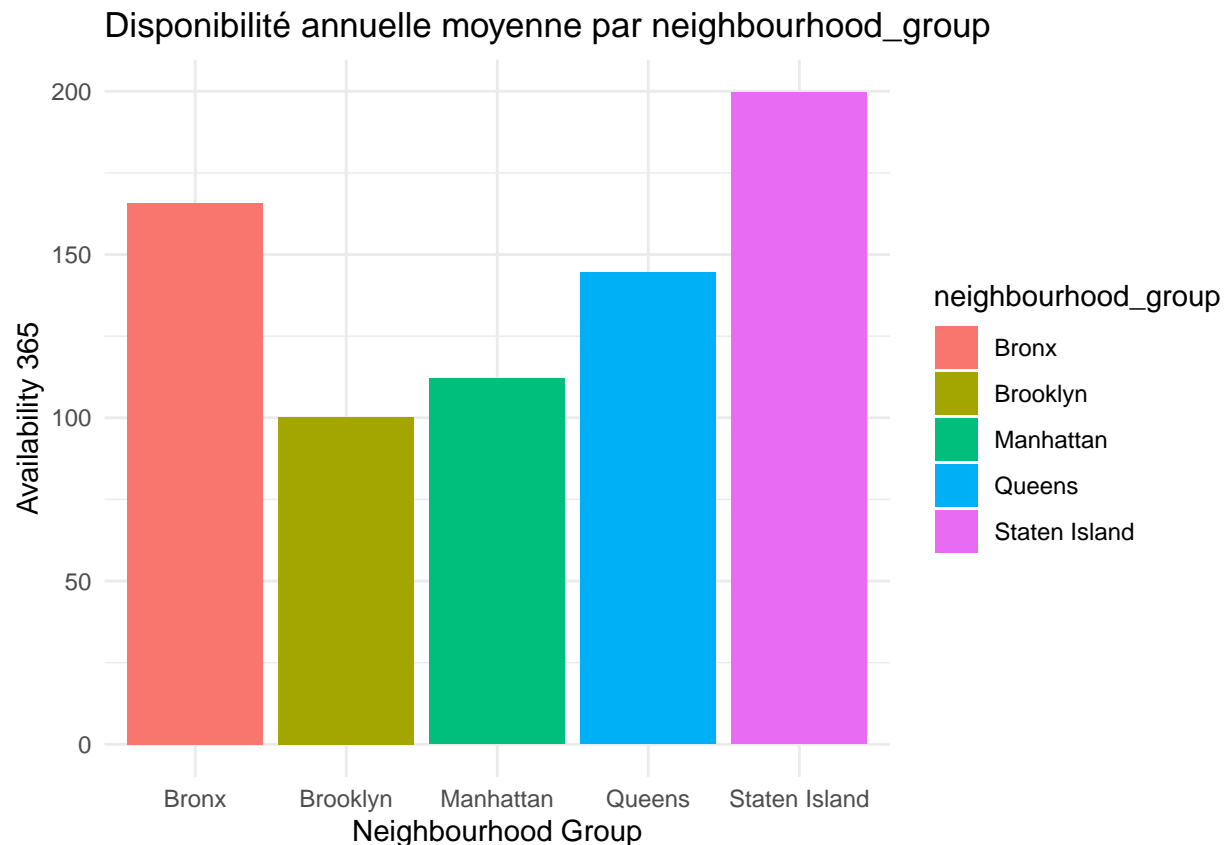
Calcul de la moyenne du prix par type de chambre

```
aggregated_data <- data %>%
  group_by(room_type) %>%
  summarize(mean_price = mean(price))
head(aggregated_data)
```

```
## # A tibble: 3 x 2
##   room_type       mean_price
##   <chr>                <dbl>
## 1 Entire home/apt      212.
## 2 Private room          89.8
## 3 Shared room           70.1
```
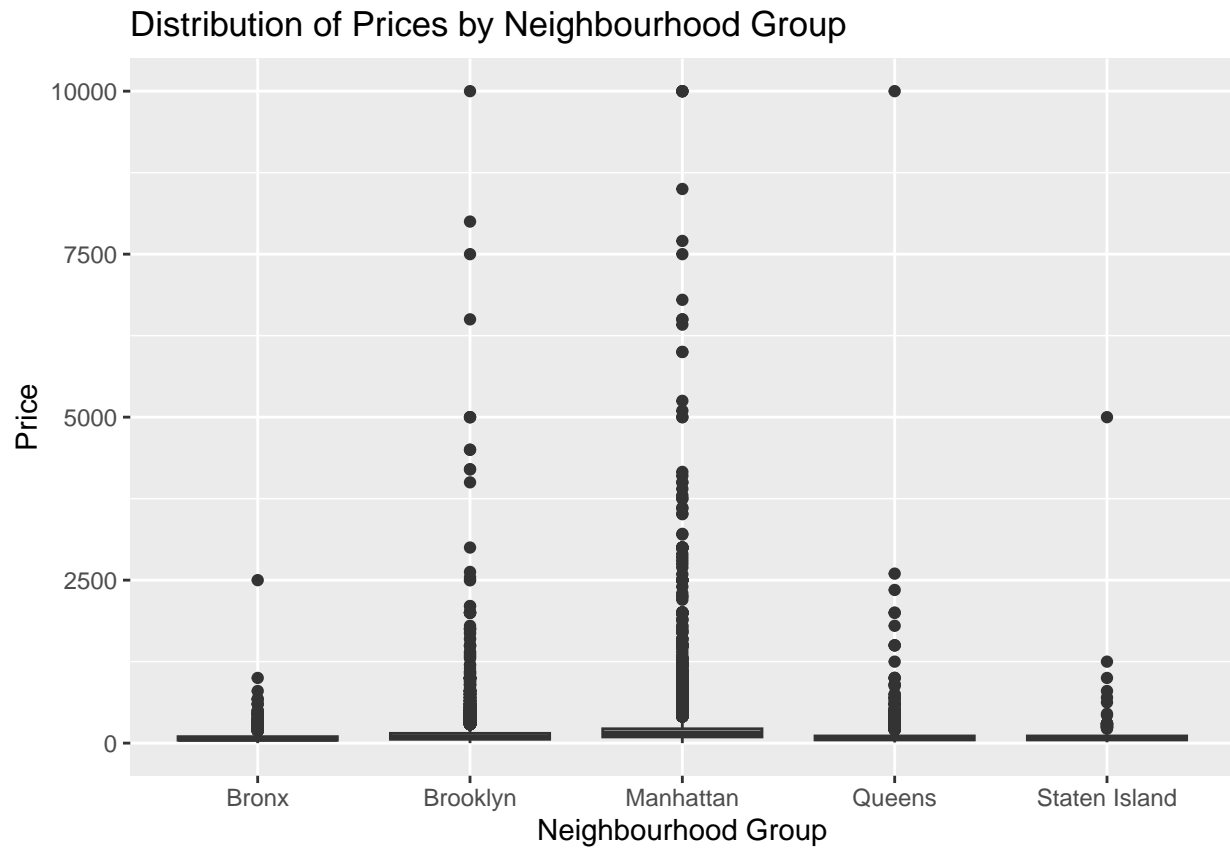
Création d'un graphique de visualisation de la disponibilité des chambres par neighbourhood

```
ggplot(data = data, aes(x = neighbourhood_group, y = availability_365, fill = neighbourhood_group)) +
  stat_summary(fun = "mean", geom = "bar") +
  labs(title = "Disponibilité annuelle moyenne par neighbourhood_group",
       x = "Neighbourhood Group",
       y = "Availability 365") +
  theme_minimal()
```
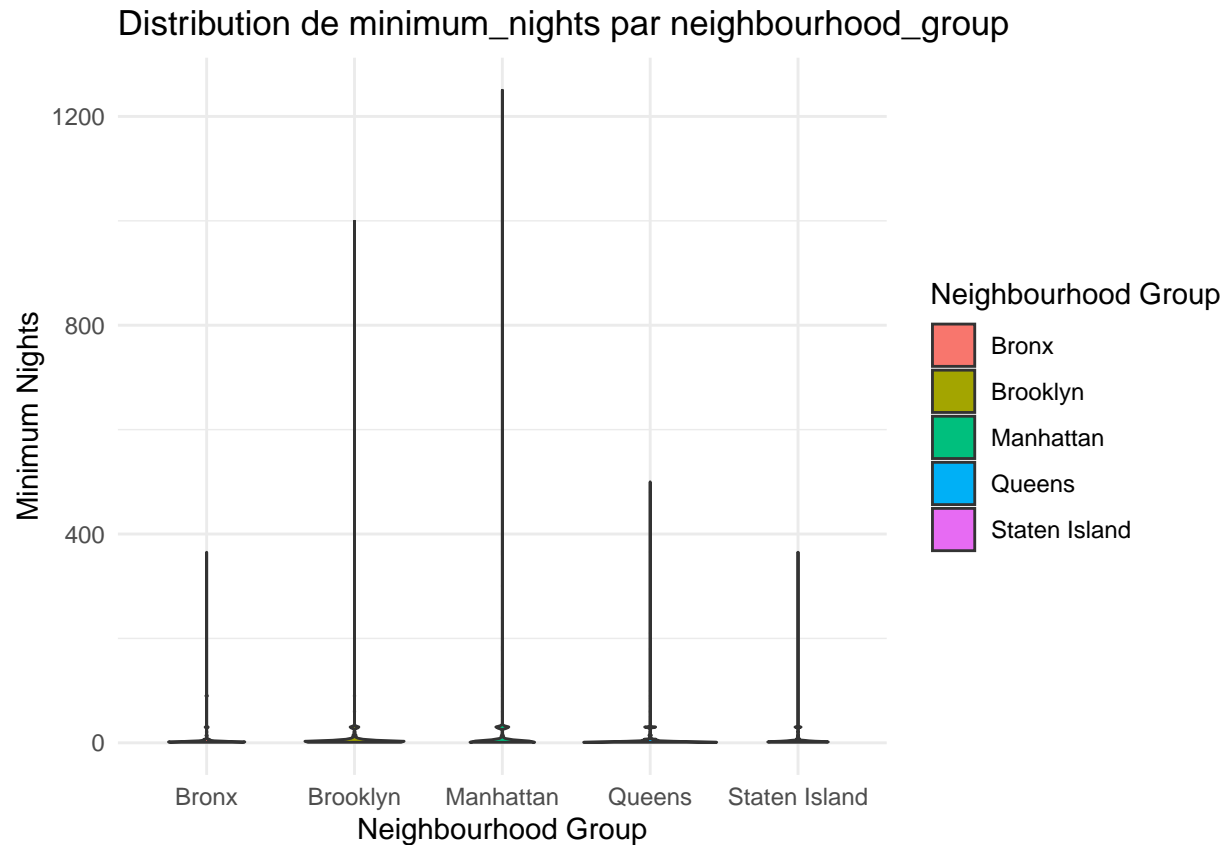
Créeation d'un graphique de visualisation des prix par neighbourhood_group

```
ggplot(data = data, aes(x = neighbourhood_group, y = price)) +
  geom_boxplot() +
  labs(x = "Neighbourhood Group", y = "Price", title = "Distribution of Prices by Neighbourhood Group")
```


Distribution of Prices by Neighbourhood Group

Créeation d'un graphique de visualisation des nuits passées par neighbourhood_group

```
ggplot(data = data, aes(x = neighbourhood_group, y = minimum_nights, fill = neighbourhood_group)) +
  geom_violin() +
  labs(title = "Distribution de minimum_nights par neighbourhood_group",
       x = "Neighbourhood Group",
       y = "Minimum Nights",
       fill = "Neighbourhood Group") +
  theme_minimal()
```

Distribution de minimum_nights par neighbourhood_group

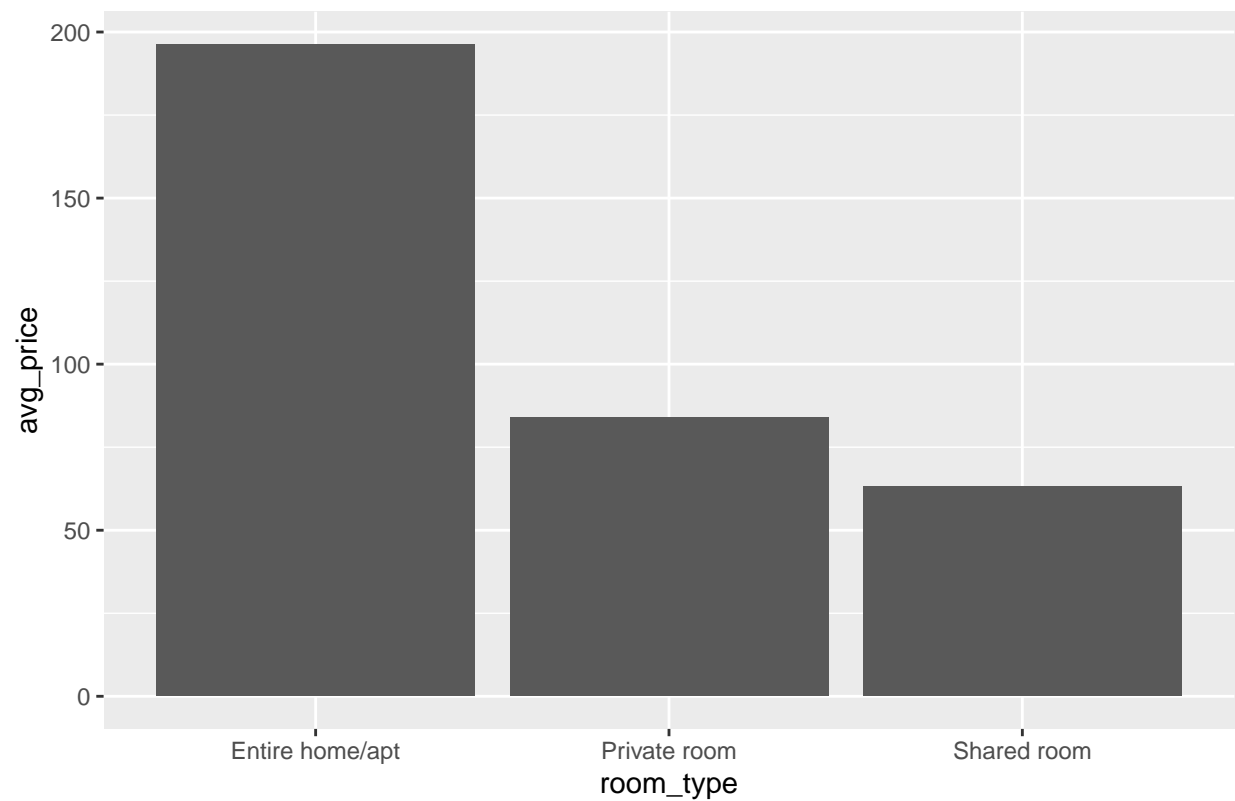Calcul de la moyenne de prix par quartier

```
summarized_by_neighbourhood <- data %>%
  group_by(neighbourhood) %>%
  summarize(mean_price = mean(price))
head(summarized_by_neighbourhood)
```

```
## # A tibble: 6 x 2
##   neighbourhood mean_price
##   <chr>              <dbl>
## 1 Allerton            87.6
## 2 Arden Heights       67.2
## 3 Arrochar           115
## 4 Arverne            172.
## 5 Astoria            117.
## 6 Bath Beach          81.8
```

Exploration des prix par type de logement

```
data_clean %>%
  group_by(room_type) %>%
  summarise(avg_price = mean(price)) %>%
  ggplot(aes(x = room_type, y = avg_price)) +
  geom_col() +  # Utilise geom_col() à la place de geom_bar()
  labs(title = "Prix moyen par type de logement")
```

Prix moyen par type de logement

Création d'une carte de location

```r
data_clean %>%
  ggplot() +
  geom_point(
    mapping = aes(x = longitude, y = latitude, color = room_type),
    size = 2
  ) +
  labs(title = "la carte des locations")
```

la carte des locations