**Examen de Statistiques**

Nom : ……………………………………………

……………………………………………

Section :………………………………………………

**1h30**

Donner la commande qui permet d’importer la base de données Hotels à partir du fichier suivant : C:\Users\ENSAM\Desktop\Hotels.csv

NB. : le fichier contient une ligne entête des colonnes, négliger les autres options de la commande

>……………………………………………………………………………………………………………………………………………………………………………………………………………………

> Hotels

PAYS ETOILE CONFORT CHAMBRE CUISINE SPORT PLAGE PRIX

Appolpon Grèce 1 4 56 2 0 8 390

Caravel Grèce 4 7 471 7 6 5 468

Christina Grèce 2 7 93 3 0 5 427

Economy Grèce 1 3 56 1 0 8 369

Eden Beach Grèce 1 4 286 3 4 7 499

Hanikian Beach Grèce 3 6 282 5 10 10 526

Marina Beach Grèce 3 6 310 7 7 10 587

Xenia Grèce 3 4 300 6 10 8 534

Agdal Maroc 4 5 146 5 1 0 447

Almohades Maroc 4 6 250 8 0 3 482

Atlas Maroc 4 5 196 9 6 6 511

Atlas Arnadi I Maroc 4 7 324 10 6 5 532

Chems Maroc 4 5 138 3 2 0 450

Dunes d'or Maroc 3 4 400 10 10 10 569

La Résidence Maroc 0 5 366 7 4 10 419

M'Diq Maroc 0 3 300 5 7 10 421

Omayades Maroc 4 6 144 7 4 8 579

Sahara Maroc 5 7 330 10 5 5 598

Sud Bahia Maroc 4 5 260 5 2 6 495

Tour Hassan Maroc 5 7 170 10 1 10 730

Alfa Mar Portugal 4 6 254 7 10 8 646

App Do Mar Portugal 4 8 140 7 6 10 652

De Lagos Portugal 4 6 273 10 10 10 802

Madeira Portugal 5 7 260 10 8 10 761

Reid's Portugal 5 7 169 10 10 10 1101

Abou Sofiane Tunisie 3 4 225 5 10 10 434

Asdrubal Tunisie 4 4 225 7 6 10 489

Colombes Tunisie 3 5 250 9 10 8 436

Darnaour Tunisie 2 3 550 6 9 8 399

Djerba Tunisie 3 6 800 10 10 10 477

Mimosa Tunisie 2 4 150 5 6 4 375

Omarit Tunisie 3 5 425 7 7 8 486

Orangers Tunisie 3 4 366 5 8 8 447

Penelope Tunisie 0 5 200 5 10 7 473

Ulysse Tunisie 4 4 130 8 7 10 495

Dauphin Or Turquie 3 4 500 8 10 10 617

Hippocampe Turquie 1 2 50 1 5 10 489

Kuriat Palace Turquie 4 9 232 10 10 10 520

Sultan Turquie 0 3 110 7 6 8 534

> ………………………………………………………%attacher la base de donner Hotels

Donner la taille de la population : ……..…

Citer deux Exemples de :

* Unités statistiques étudiés :

…………………………………..…………

………………………………..……………

* Caractère statistique étudiés :

………………………………..……………

………………………………………………

> ………………………………………………………%afficher la structure de Hotels

'data.frame': 39 obs. of 8 variables:

$ PAYS : chr "Grèce" "Grèce" "Grèce" "Grèce" ...

$ ETOILE : int 1 4 2 1 1 3 3 3 4 4 ...

$ CONFORT: int 4 7 7 3 4 6 6 4 5 6 ...

$ CHAMBRE: int 56 471 93 56 286 282 310 300 146 250 ...

$ CUISINE: int 2 7 3 1 3 5 7 6 5 8 ...

$ SPORT : int 0 6 0 0 4 10 7 10 1 0 ...

$ PLAGE : int 8 5 5 8 7 10 10 8 0 3 ...

$ PRIX : int 390 468 427 369 499 526 587 534 447 482 ...

>

> ………………………………………………………%lister les 3 premières lignes de Hotels

PAYS ETOILE CONFORT CHAMBRE CUISINE SPORT PLAGE PRIX

Appolpon Grèce 1 4 56 2 0 8 390

Caravel Grèce 4 7 471 7 6 5 468

Christina Grèce 2 7 93 3 0 5 427

> ………………………………………………………………… % afficher la colonne ETOILE

[1] 1 4 2 1 1 3 3 3 4 4 4 4 4 3 0 0 4 5 4 5 4 4 4 5 5 3 4 3 2 3 2 3 3 0 4 3 1 4 0

> ………………………………………………………%extraire les modalités de ETOILE

[1] 1 4 2 3 0 5

> …………………………………………………………% donner la table des effectifs de ETOILE

ETOILE

0 1 2 3 4 5

4 4 3 10 14 4

> ……………………………………………………………% dcpr (c’est-à-dire : **d**éduire la **c**ommande à **p**artir du **r**ésultat)

0 1 2 3 4 5

4 8 11 21 35 39

|  |  |  |
| --- | --- | --- |
| > ……………………………………………………………………% dcpr | > ………………………………………………………% dcpr | > …………………………………………………………% dcpr |
|  |  |  |

Méthode de calcul

Déduire les valeurs de :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| mode | min | max | median | Intervalle inter quartile | L’étendue |
|  |  |  |  |  |  |

> …………………………………………………………………………% dcpr et compléter le résultat

Min. 1st Qu. Median Mean 3rd Qu. Max.

…………… …………… ……………… …………… ………………… ………………

NB. Dans cette dernière question, il y a une valeur que vous devez calculer (ici 🡪 ) à partir des résultats des questions précédentes, les autres sont à déduire.

> …………………………………………………………………………………………………………………………………………………………………………………%la variance théorique de ETOILE

[1] 2.17883

> …………………………………………………………………………………………………………………………………………………………………………………%la variance corrigée de ETOILE

[1] 2.236167

> ……………………………………………………………% nb des hôtels par payer

PAYS

Grèce Maroc Portugal Tunisie Turquie

8 12 5 10 4

> ………………………………………………………………………………………………………………………………………………… %les moyennes conditionnées/PAYS

Group.1 ETOILE CONFORT CHAMBRE CUISINE SPORT PLAGE PRIX

Calculer ici la valeur manquante

1 Grèce 2.250000 5.125000 231.75 4.250000 4.625 7.625000 475.0000

2 Maroc 3.416667 5.416667 252.00 7.416667 4.000 6.083333 519.4167

3 Portugal 4.400000 6.800000 219.20 8.800000 8.800 9.600000 792.4000

4 Tunisie 2.700000 4.400000 332.10 6.700000 8.300 8.300000 451.1000

5 Turquie 2.000000 4.500000 223.00 6.500000 7.750 9.500000 540.0000

> …………………………………………………………………………………………………………………………………………………………………………% variance conditionnées /PAYS

Group.1 ETOILE CONFORT CHAMBRE CUISINE SPORT PLAGE PRIX

1 Grèce 1.357143 2.4107143 22073.929 5.357143 18.553571 3.696429 5716.571

2 Maroc 2.810606 1.5378788 8572.364 6.083333 8.727273 13.719697 8053.356

3 Portugal 0.300000 0.7000000 3640.700 2.700000 3.200000 0.800000 34364.300

4 Tunisie 1.344444 0.7111111 44147.433 3.344444 2.900000 3.566667 1641.656

5 Turquie 3.333333 9.6666667 39836.000 15.000000 6.916667 1.000000 2988.667

Calculer la variance marginale de ETOILE

Nous allons maintenant chercher le modèle de regression multiple de PRIX en ETOILE , CONFORT, CHAMBRE, CUISINE, SPORT et PLAGE

> reg\_mult = ……………(PRIX ~ ETOILE+CONFORT+CHAMBRE+CUISINE+SPORT+PLAGE)%compléter cette commande

> …………………………………………………………………………………………………………………………………………………………………………% dcpr

Call:

lm(formula = PRIX ~ ETOILE + CONFORT + CHAMBRE + CUISINE + SPORT + PLAGE)

Residuals:

Min 1Q Median 3Q Max

-199.70 -25.95 -8.20 35.72 355.98

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 200.3860 77.1080 2.599 0.0140 \*

ETOILE 31.2705 15.8378 1.974 0.0570 .

CONFORT 12.0172 14.1751 0.848 0.4029

CHAMBRE -0.2871 0.1303 -2.204 0.0348 \*

CUISINE 14.5403 10.1264 1.436 0.1607

SPORT 5.6717 6.5173 0.870 0.3906

PLAGE 15.0559 7.3588 2.046 0.0491 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 99.41 on 32 degrees of freedom

Multiple R-squared: 0.5622, Adjusted R-squared: 0.4801

F-statistic: 6.848 on 6 and 32 DF, p-value: 9.66e-05

Selon ce modèle de régression linéaire donner l’expression mathématique qui permet de calculer le PRIX en fonction des autres variables explicatives

…………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………….

> …………………………………………………% donner les valeurs (PRIX) correspondent aux prédictions du modèle de régression

1 2 3 4 5 6 7 8 9 10 11

413.1765 485.4625 439.2490 386.6189 369.3176 565.3199 569.3470 520.5463 422.0126 487.2899 584.5135

12 13 14 15 16 17 18 19 20 21 22

571.2853 400.9003 580.1108 430.4265 413.2742 601.1471 595.1615 485.2919 693.6883 603.5976 667.7851

23 24 25 26 27 28 29 30 31 32 33

671.8758 707.5522 745.0205 557.6494 595.3137 590.5389 399.8152 489.3105 434.8882 494.2030 475.7149

34 35 36 37 38 39

437.8647 642.7990 522.3214 434.7945 719.6980 461.1179

Ecrire le code qui permet de prévoir le prix pour un hôtel avec la configuration suivante :

ETOILE = 4 ; CONFORT = 7 ; CHAMBRE = 90 ; CUISINE = 3 ; SPORT = 0 ; PLAGE 400

Nous allons maintenant retirer les variables non significatives du model précédant, proposer la (les) variable(s) qu’il faut retirer : …………………………………………………………………………………………………………………............................

Justifier votre réponse : …………………………………………………………………………………………………………..

……………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………