Rapport Phase 3 SAé BDD:

Groupe: Cygnus

Gupta

Varun Vedic

Bousserhane

Salma

I-Création des tables :

Nous avons commencé à créer précisément 3 tables (relations).

a)1ère Table Passagers

Dans la première Table que nous avons nommée"Passagers", nous classons dans plusieurs attributs les informations du passager lors de la croisière du Titanic.

```
create table Passagers(
Name varchar(200) not null,
Sex varchar(10),
Age float,
SibSp int,
Parch int,
primary key(Name)
```

b) 2ème Table Trajet

Dans la deuxième table que nous avons nommée "Trajet" où sont classées les données portant sur le séjour des passagers en lui-même.

```
create table Trajet(
    PassengerId int not null,
    Pclass int,
    Ticket varchar(50),
    Fare float,
    Cabin varchar(20),
    primary key(PassengerId) );
```

c) 3ème Table titanic

Pour finir, la dernière relation que nous avons nommée "titanic" où sont classées les informations portant sur les embarcations et les survivants après la collision du Titanic.

```
create table titanic(
    Name varchar(200) not null,
    PassengerId int not null,
    Survived int,
    Enbarked varchar(5),
    foreign key(Name) references Passagers(Name),
    foreign key(PassengerId) references Trajet(PassengerId) );
```

II- Insérer des données dans les différentes tables :

Nous avons ensuite à partir du fichier csv où sont contenus toutes les données, et aussi à partir d'excel trier et sélectionner toutes les données en fonction des tables. Nous avons donc créé différents fichiers .txt correspondant aux données de chaque table.

```
\copy passagers from 'C:\Users\visha\Documents\passagers.txt'
\copy trajet from 'C:\Users\visha\Documents\trajet.txt'
\copy titanic from 'C:\Users\visha\Documents\titanic.txt'
```

a)Visualisation de la table Passagers

titanic=> select * from Passagers; name	sex	age	sibsp	parch
			· 	+
Braund, Mr. Owen Harris	male	22	1	0
Cumings, Mrs. John Bradley (Florence Briggs Thayer)	female	38	1	0
Heikkinen, Miss. Laina	female	26	0	0
Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35	1	0
Allen, Mr. William Henry	male	35	0	0
Moran, Mr. James	male	0	0	9
McCarthy, Mr. Timothy J	male	54	0	0
Palsson, Master. Gosta Leonard	male	2	3	1
Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27	0	2
Nasser, Mrs. Nicholas (Adele Achem)	female	14	1	0
Sandstrom, Miss. Marguerite Rut	female	4	1	1
Bonnell, Miss. Elizabeth	female	58	0	0
Saundercock, Mr. William Henry	male	20	0	0
Andersson, Mr. Anders Johan	male	39	1	5
Vestrom, Miss. Hulda Amanda Adolfina	female	14	0	0
Hewlett, Mrs. (Mary D Kingcome)	female	55	0	0
Rice, Master. Eugene	male	2	4	1
Williams, Mr. Charles Eugene	male	0	0	0

b)Visualisation de la table Trajet

D, Froduitouriori				
titanic=> sele	ect * from	ı Trajet;		
passengerid	pclass	ticket	fare	cabin
	+	+	+	+
1	3	A/5 21171	7.25	l
2	1	PC 17599	71.2833	C85
3	3	STON/02. 3101282	7.925	İ
4	1	113803	53.1	C123
5	3	373450	8.05	l
6	3	330877	8.4583	l
7	1	17463	51.8625	E46
8	3	349909	21.075	l
9	3	347742	11.1333	l
10	2	237736	30.0708	
11	3	PP 9549	16.7	G6
12	1	113783	26.55	C103
13	3	A/5. 2151	8.05	

c)Visualisation de la table titanic

titanic=> select * from titanic;			
name	passengerid	survived	enbarked
Braund, Mr. Owen Harris	1	0	S
Cumings, Mrs. John Bradley (Florence Briggs Thayer)	2	1	С
Heikkinen, Miss. Laina	3	1	S
Futrelle, Mrs. Jacques Heath (Lily May Peel)	4	1	S
Allen, Mr. William Henry	5	0	S
Moran, Mr. James	6	0	Q
McCarthy, Mr. Timothy J	7	0	S
Palsson, Master. Gosta Leonard	8	0	S
Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	9	1	S
Nasser, Mrs. Nicholas (Adele Achem)	10	1	C

III-Requêtes

a) Requêtes imaginées

-- lère requête : Quels sont les noms et le sexe des passagers qui ont un âge inférieur à 40 ans select name, sex from Passagers where age<40;</p>

name	sex
Braund, Mr. Owen Harris Cumings, Mrs. John Bradley (Florence Briggs Thayer)	male female
Heikkinen, Miss. Laina Futrelle, Mrs. Jacques Heath (Lily May Peel)	female female
Allen, Mr. William Henry	male
Moran, Mr. James	male
Palsson, Master. Gosta Leonard Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	male female
Nasser, Mrs. Nicholas (Adele Achem)	female
Sandstrom, Miss. Marguerite Rut Saundercock, Mr. William Henry	female male
Andersson, Mr. Anders Johan	male

Dans la **première requête**, nous pouvons voir qu'environ **728 passagers(voyageurs)** ont un **âge en dessous de 40 ans**. Ce qui veut dire que sur un total de passagers qui est égal à **891 passagers** (on peut le voir grâce à la requête n°4), les passagers d'en dessous de 40 ans représentent 81,7 % de tous les passagers.

```
--2ème requête : Même Requête que la 1 et qui ont pour sexe, le sexe masculin
select name,sex
from Passagers
where age<40 and sex='male';
```

name	sex
Braund, Mr. Owen Harris Allen, Mr. William Henry Moran, Mr. James Palsson, Master. Gosta Leonard Saundercock, Mr. William Henry Andersson, Mr. Anders Johan	male male male male male male
Rice, Master. Eugene	male

Dans la **deuxième requête**, nous filtrons encore plus la demande en y indiquant qu'on veut seulement ceux appartenant au genre masculin, nous obtenons un résultat de **468 passagers** du genre masculin, ce qui représente **64,28** % sur les

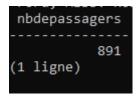
728 passagers, on peut donc conjecturer qu'il y'a plus d'hommes que de femmes lors de la croisière du Titanic.

--3 ème Requête: Donner le nom,l'age des passagers qui ont un prix de ticket supérieur à 20 et qui ont embarqués depuis le port S. select Name,age from Passagers natural join titanic natural join Trajet where fare>20 and enbarked='S';

name	age .
Futrelle, Mrs. Jacques Heath (Lily May Peel)	35
McCarthy, Mr. Timothy J	54
Palsson, Master. Gosta Leonard	2
Bonnell, Miss. Elizabeth	58
Andersson, Mr. Anders Johan	39
Fynney, Mr. Joseph J	35
Sloper, Mr. William Thompson	28
Palsson, Miss. Torborg Danira	8
Asplund, Mrs. Carl Oscar (Selma Augusta Emilia Johansson)	38
Fortune, Mr. Charles Alexander	19

Dans la troisième requête, nous avons essayé de faire une requête en utilisant les jointures naturelles pour relier les trois tables (relations), nous interrogeons les tables sur une demande bien précise, nous demandons les noms et âge des passagers selon différents critères que le prix du ticket est supérieur à 20 et qu'ils ont bien embarqué depuis le port S (Southampton), ils représentent 264 passagers sur 891 soit environ 30%.

```
-- 4 ème requête Donnez le nombre de Passagers
select count(*) as NbdePassagers from Passagers;
```



Dans la quatrième requête, on obtient le nombre total de passagers qui est égal à 891 passagers.

```
-- Sème Requête: Donnez le prix moyen pour la croisière du Titanic
select avg(Fare) as PrixMoyenTicket from Trajet;
```

Dans la **cinquième requête**, nous utilisons une fonction d'agrégation **avg(fare)** qui va calculer le prix moyen d'un ticket, qui va donner environ 32,4 qui est donc le prix moyen d'un ticket de passagers lors de la croisière du Titanic.

```
--6ème requête : Donner longueur (nb de caractères) du prénom le plus long select max(char_length(Name)) as longueur_chaine_de_caractère from Passagers;
--Pour savoir le nom on a juste à faire:
select Name from Passagers where char_length(Name)=82;
```

```
longueur_chaine_de_caract | re

82

(1 ligne)

name

Penasco y Castellana, Mrs. Victor de Satode (Maria Josefa Perez de Soto y Vallejo)
(1 ligne)
```

Dans la **sixième requête**, la longueur du prénom le plus grand grâce à la fonction **d'agrégation** max(char_length(attribut)).

```
--7ème requête:Combien de passagers y'avait t-il depuis chaque embarcation ?
select enbarked,count(*) as NbdeVoyageurs
from titanic
group by enbarked
order by enbarked;
```

Dans la **septième requête**, nous cherchons combien de passagers ont embarqué depuis le **port DE(Southampton)**, depuis le **port C (Cherbourg) et depuis le port**

Q(Queenstown). 2 passagers n'ont aucune embarcation assignée,168 sont assignés aux embarcations depuis le port C,quant à Q 77 passagers et pour finir 644 passagers qui ont fait embarcation depuis le port S.Cela représente O.22%(passagers ayant aucune embarcation) sur les 891 passagers. L'embarcation des passagers depuis le port C représente environ 19% sur les 891 passagers.L'embarcation des passagers depuis le port Q représente 8,64% sur les 891 passagers .L'embarcation des passagers depuis le port S représente 72.27% sur les 891 passagers. On observe qu'il y'a eu majoritairement des embarcations depuis le port S, ce qui est normal vu que c'est la 1ère escale du Titanic.

```
--8ème requête: Donner le nombre de passagers morts lors de la croisière du Titanic. select count(*) as nbPassagersmorts from titanic where survived=0;
```

```
nbpassagersmorts
-----
549
(1 ligne)
```

Dans la **huitième requête**, nous voyons grâce à la fonction d'agrégation count(), le nombre de morts lors du trajet du Titanic est de **549**, ce qui représente plus de la moitié des passagers morts(61% environ) sur les 891 passagers.

b)Requêtes imposées

-- lère requête:Combien de classes de passagers différentes y'avait-il à bord du titanic? select count(distinct Pclass) as Classes_differentes from trajet;

```
classes_differentes
-----3
(1 ligne)
```

Dans la **première requête**, nous demandons combien de classes différentes étaient disponibles, on peut voir qu'il y'a 3 classes différentes à bord du titanic, **la première** classe 1 était destinées aux gens les plus fortunés du navire.La deuxième classe 2 était destinée aux gens modestes (enseignants) et la dernière classe était destinée aux gens de la classe populaire ou venaient de l'immigration.

```
--2ème requête:Combien de passagers y'avait t-il dans chaque classe ?
select Pclass, count(*) as NbDeVoyageursDansClasse
from Trajet
group by Pclass
order by Pclass;
```

pclass	nbdevoyageursdansclasse
1	216
2	184
3	491
(3 lignes)

Dans la deuxième requête, nous nous demandons le nombres de passagers par classe, dans la 1ère classe, il y'a 216 passagers soit 25 % sur la totalité des passagers, dans la seconde classe, il y'a 184 passagers soit 20.6 % sur la totalité des passagers et enfin pour finir dans la dernière classe, il y'a 491 passagers soit 55.10 % sur la totalité des passagers. Il y'a deux fois plus de passagers dans la classe 3 que dans la classe 1, cela peut être expliqué par le fait que beaucoup de gens à cette époque voulaient immigrer pour une meilleure vie.

```
--3ème requête:Combien de femmes et d'hommes y'avait-il dans chaque classe ?
select Pclass, sex, count(*) as VoyageursClasse
from Passagers natural join titanic natural join Trajet
group by Pclass,sex
order by Pclass;
```

pclass	sex	voyageursclasse
1 1 2 2 3 3 (6 lignes	female male male female female female	94 92 122 108 76 144 347

Dans la troisième requête, nous voulons savoir le nombre d'hommes et de femmes dans chaque classe. Dans la 1ère classe, il y'a 94 passagers femmes et 122 passagers hommes donc 216 passagers au total, soit 43.55 % des passagers sont des femmes dans la classe 1 et donc 56.4 % les passagers hommes. Dans la 2ème classe, il y'a 76 passagers femmes et 108 passagers hommes donc 184 passagers au total, soit 41.30% des passagers sont des femmes dans la classe 2 et donc 58.6 % les passagers hommes. Dans la 3ème classe, il y'a 144 passagers femmes et 347 passagers hommes donc 491 passagers au total, soit 29.3% des

passagers sont des femmes dans la classe 3 et donc **70.6** % les passagers hommes.On peut observer que dans les classes 1 et 2 qu'il y'a à peu près le même pourcentage d'hommes et de femmes.Cependant on peut observer un pourcentage nettement plus élevé pour les hommes par rapport aux femmes (**environ 70**% **pour H et 30**% **pour F)**

```
--4 ème requête: Compter le nombre et le pourcentage de survivants et de passagers morts.

select Survived, count(*) Voyageurs, (count(*)*100/(select count(*) from titanic)) as pourcent from titanic group by survived;
```

survived	voyageurs	pourcent
0 1 (2 lignes)	549 342	61 38

Dans la **quatrième requête**, on obtient le résultat suivant **549 voyageurs** qui n'ont pas survécu soit **549/891x100=61.1%** de voyageurs morts et **342 voyageurs qui ont survécu soit 342/891x100=38.38%**.On peut voir un pourcentage de morts plus élevé que le pourcentage des voyageurs qui ont survécu.

```
--5ème requête: Visualiser la répartition des passagers survivants et morts par classe create view survivants_selon_classe (ClasseDesPassagers,Survivants,Passagers) as select Pclass,survived, count(*) as Voyageurs from Trajet natural join titanic group by Pclass,survived order by Pclass,survived;
```

select * from survivants_selon_classe;

CREATE VIEW classedespassagers	survivants	passagers
1	0	80
1	1	136
2	0	97
2	1	87
3	0	372
3	1	119
(6 lignes)		

Dans la cinquième requête, nous filtrons la requête en comptant le nombre de passagers vivants et morts en fonction de la classe. Dans la classe 1, il y'a 80 passagers qui n'ont pas survécu et 136 passagers qui ont survécu, on peut voir qu'il y'a eu plus de survivants que de morts. Dans la classe 2, il y'a 97 passagers

qui n'ont pas survécu et 87 passagers qui ont survécu, on peut voir qu'il y'a environ le même nombre de morts que de survivants. Dans la classe 3, il y'a 372 passagers qui n'ont pas survécu et 119 passagers qui ont survécu, on peut observer qu'il y'a eu énormément plus de morts que de survivants. On peut aussi apercevoir qu'il y'a eu plus de sauvés dans la classe 1 que dans la classe 3 sachant qu'il y'a beaucoup plus de passagers en classe 3 qu'en classe 1, on peut apercevoir qu'il y'a eu une priorité de sauvetage, cela est aussi peut être dû au faite que la classe 3 se situait dans les parties inférieures du bateau et la classe 1 dans les parties supérieures du bateau.

```
--6ème requête: Visualiser la répartition des passagers survivants et des passagers décédés selon sexe. create view survivants_selon_genre (Survivants,GenreDuPassager,NombreDePassagers) as select survived,sex, count(*) as Voyageurs from Passagers natural join titanic group by survived,sex order by survived,sex; select * from survivants_selon_genre;
```

```
CREATE VIEW
survivants | genredupassager | nombredepassagers

0 | female | 81
0 | male | 468
1 | female | 233
1 | male | 109
(4 lignes)
```

Dans la dernière requête, on filtre le nombre de survivants et de morts en fonction du sexe. On obtient le résultat suivant:

- -Pour les **femmes**, il y'a **81 passagers** qui n'ont pas survécu et **233 qui ont survécu**. Il y'a eu plus de survivants que de morts.
- -Pour les **hommes**, il y'a eu **468 passagers** qui n'ont pas survécu et **109 qui ont survécu**. Il y'a eu beaucoup plus de morts que de survivants.

IV-Annexes:

a)Visualisation des fichiers .txt où se trouve les données.

passagers.txt - Bloc-notes

Fichier Edition Format Affichage Aide							
Braund, Mr. Owen Harris male 22	2	1	0				
Cumings, Mrs. John Bradley (Florer	nce Bri	iggs Tha	yer)	female	38	1	0
Heikkinen, Miss. Laina female 20	.6	0	0				
Futrelle, Mrs. Jacques Heath (Lily	y May F	Peel)	female	35	1	0	
Allen, Mr. William Henry ma	ale	35	0	0			
Moran, Mr. James male 0)	0	0				
McCarthy, Mr. Timothy J male 54	4	0	0				
Palsson, Master. Gosta Leonard ma	ale	2	3	1			
Johnson, Mrs. Oscar W (Elisabeth \	Vilhelm	mina Ber	g)	female	27	0	2
Nasser, Mrs. Nicholas (Adele Achem	em)	female	14	1	0		
Sandstrom, Miss. Marguerite Rut fe	emale	4	1	1			
Bonnell, Miss. Elizabeth fe	emale	58	0	0			
Saundercock, Mr. William Henry ma	ale	20	0	0			
Andersson, Mr. Anders Johan ma	ale	39	1	5			
Vestrom, Miss. Hulda Amanda Adolfi	ina	female	14	0	0		

trajet.txt - Bloc-notes

Fichier	Edition	Format Affichage Aide
1	3	A/5 21171 7.25
2	1	PC 17599 71.2833 C85
3	3	STON/02. 3101282 7.925
4	1	113803 53.1 C123
5	3	373450 8.05
6	3	330877 8.4583
7	1	17463 51.8625 E46
8	3	349909 21.075
9	3	347742 11.1333
10	2	237736 30.0708
11	3	PP 9549 16.7 G6
12	1	113783 26.55 C103
13	3	A/5. 2151 8.05
14	3	347082 31.275
15	3	350406 7.8542
16	2	248706 16
17	3	382652 29.125
18	2	244373 13

titanic.txt - Bloc-notes					
Fichier Edition Format Affichage Aide					
Braund, Mr. Owen Harris 1 0	S				
Cumings, Mrs. John Bradley (Florence	e Briggs T	nayer)	2	1	C
Heikkinen, Miss. Laina 3 1	S				
Futrelle, Mrs. Jacques Heath (Lily M	May Peel)	4	1	S	
Allen, Mr. William Henry 5	0	S			
Moran, Mr. James 6 0	Q				
McCarthy, Mr. Timothy J 7 0	S				
Palsson, Master. Gosta Leonard 8	0	S			
Johnson, Mrs. Oscar W (Elisabeth Vil	lhelmina B	erg)	9	1	S
Nasser, Mrs. Nicholas (Adele Achem)	10	1	C		
Sandstrom, Miss. Marguerite Rut 11	1	S			
Bonnell, Miss. Elizabeth 12	1	S			
Saundercock, Mr. William Henry 13	0	S			
Andersson, Mr. Anders Johan 14	0	S			
Vestrom, Miss. Hulda Amanda Adolfina	a 15	0	S		
Hewlett, Mrs. (Mary D Kingcome)	16	1	S		
Rice, Master. Eugene 17 0	Q				

Script et Rapport en mode Notebook

Fonotions d'agrégations

4) Donnez le nombre de Passagers

```
Entrée [7]: H 1 XXsql 2 select count(*) as NbdePassagers
                        from Passagers;
```

Environment variable \$DATABASE_URL not set, and no connect string given. Connection info needed in SQLAIchemy format, example: postgresql://username:password@hostname/dbname or an existing connection: dict_keys([])

5) Donnez le prix moyen pour la croisière du Titanic.

```
Entrée [8]: M 1 %%sql
                   2 select avg(Fare)
3 from Trajet;
```

Environment variable \$DATABASE_URL not set, and no connect string given.

Connection info needed in SQLAIChemy format, example:

postgresql://username:password@hostname/dbname
or an existing connection: dict_keys([])

6) Donner longueur (nb de caractères) du prênom le plus long

```
Entrée [9]: H 1 XXsql 2 select max(char_length(Name)) 3 from Passagers;
```

7) Combien de passagers y'avait t-il depuis chaque embarcation ?

```
Entrée [10]: M 1 XXsql 2 select enbarked,count(*) as NbdeVoyageurs 3 from titanic 4 group by enbarked 5 order by enbarked;
```

Environment variable \$DATABASE URL not set, and no connect string given.

Connection info needed in SQLAIchemy format, example:

postgresql://username:password@hostname/dbname
or an existing connection: dict_keys([])

8) Donner le nombre de passagers morts lors de la croisière du Titanic.

```
Entrée [11]: M 1 %%sql 2 select count(*) as mbPassagersmorts
                     3 from titanic
4 where survived=0;
```

Environment variable \$DATABASE_URL not set, and no connect string given.

Connection info needed in SQLAIchemy format, example:

postgresql://username:password@hostname/dbname
or an existing connection: dict_keys([])

Réponse aux requêtes de la SAé

1. Combien de classes de passagers différentes y avait-il à bord du Titanic ?

Environment variable \$DATABASE_URL not set, and no connect string given.

Connection info needed in SQLAIchemy format, example:

postgresql://username:password@hostname/dbname

or an existing connection: dict_keys([])

2. Combien de passagers y avait-il dans chaque classe?

Environment variable \$DATABASE_URL not set, and no connect string given.

Connection info needed in SQLAIchemy format, example:
postgresql://username:password@hostname/dbname
or an existing connection: dict_keys([])

3. Combien de femmes et d'hommes y avait-il dans chaque classe?

```
Entrée [14]: H

1 XXsql
2 select Pclass, sex, count(*) as VoyageursClasse
4 from Passagers natural join titanic natural join Trajet
5 group by Pclass, sex
6 order by Pclass;
```

Environment variable \$DATABASE_URL not set, and no connect string given.

Connection info needed in SQLAIchemy format, example:

postgresql://username:password@hostname/dbname

or an existing connection: dict_keys([])

4. Comptez le nombre et le pourcentage de survivants et de passagers morts.

Environment variable \$DATABASE_URL not set, and no connect string given.

Connection info needed in SQLAIChemy format, example:
postgresql://username:password@hostname/dbname
or an existing connection: dict_keys([])

5. Visualiser la répartition des passagers survivants et morts par classe.

Environment variable \$DATABASE_URL not set, and no connect string given.

Connection info needed in SQLAIChemy format, example:

postgresql://username:password@hostname/dbname

or an existing connection: dict_keys([])

6. Visualiser la répartition des passagers survivants et des passagers décédés selon le sexe.

Environment variable \$DATABASE_URL not set, and no connect string given.

Connection info needed in SQLAlchemy format, example:

postgresql://username:password@hostname/dbname

or an existing connection: dict_keys([])