



Ζήτημα 1^ο

1.

a. Δημιουργία βάσης δεδομένων με όνομα CTDW :

- CTDW
 - Database Diagrams
 - Tables
 - Views
 - External Resources
 - Synonyms
 - Programmability
 - Query Store
 - Service Broker
 - Storage
 - Security

b. Δημιουργία πίνακα CardsTransactions :

```
create table CardsTransactions(  
pid int,  
pname varchar(50),  
age int,  
gender char(1),  
cardno char(16),  
card_brand varchar(30),  
card_type varchar(20),  
tdate datetime,  
amount decimal(6,2),  
ttc int,  
trans_type varchar(30),  
mcc int,  
merchant_city varchar(50)  
);
```

c. Φόρτωση δεδομένων στον πίνακα CardsTransactions:

```
BULK INSERT CardsTransactions  
FROM  
'C:\Users\alvio\Downloads\CardsTransactions\CardsTransactions.  
txt'  
WITH (FIRSTROW =2, FIELDTERMINATOR='|', ROWTERMINATOR = '\n');
```

2. Σύμφωνα με την περιγραφή, εντοπίζονται οι εξής διαστάσεις (*dimensions*) και μετρήσιμα (*measures*):

“

Οι απαιτήσεις της διοίκησης εστιάζουν στην ανάλυση του αριθμού και της *αξίας* των συναλλαγών, βάσει της επωνυμίας και του είδους των *καρτών*, το φύλο και την ηλικία των *κατόχων* τους, τον *τύπο της συναλλαγών*, την *πόλη* στην οποία έλαβαν χώρα καθώς και οποιονδήποτε μεταξύ τους συνδυασμό. Εξυπακούεται ότι στην ανάλυση των δεδομένων θα πρέπει να ληφθεί υπόψη και ο παράγοντας του *χρόνου* έτσι ώστε, η διοίκηση του οργανισμού να είναι σε θέση να παράγει στατιστικές αναφορές με τα στοιχεία των συναλλαγών ανά μήνα, τρίμηνο και έτος.

”

Συνεπώς το star schema θα περιλαμβάνει έναν πίνακα για κάθε διάσταση (dimension). Η δημιουργία των πινάκων αυτών παρατίθεται παρακάτω:

```
create table cardinfo(  
cardno char(16) primary key,  
card_brand varchar(30),  
card_type varchar(20)  
);
```

...βάσει της *επωνυμίας* και του *είδους* των καρτών...

```
create table cardowner(  
pid int primary key,  
pname varchar(50),  
age int,  
gender char(1)  
);
```

...το *φύλο* και την *ηλικία* των κατόχων του ...

(συμπεριλαμβάνεται και το όνομα για λόγους εμφάνισης)

```
create table transactiontype(  
ttc int primary key,  
trans_type varchar(30)  
);
```

...τον *τύπο της συναλλαγών*...

```
create table city(  
mcc int primary key,  
merchant_city varchar(50)  
);
```

...την *πόλη* στην οποία έλαβαν χώρα...

```
create table timeinfo(  
time_key datetime primary key,  
t_year int,  
t_quarter int,  
t_month int,  
);
```

...παράγοντας του *χρόνου* έτσι ώστε, η διοίκηση του οργανισμού να είναι σε θέση να παράγει στατιστικές αναφορές με τα στοιχεία των συναλλαγών ανά *μήνα*, *τρίμηνο* και *έτος*.

Τέλος, δημιουργείται ο πίνακας που περιέχει τα γεγονότα (facts). Ένα γεγονός αντιστοιχεί σε μια συναλλαγή. Ο fact table περιέχει τα κλειδιά των dimension tables ως ξένα κλειδιά (foreign keys), που είναι και ο λόγος για τον οποίο έπρεπε να δημιουργηθεί μετά από τα

dimension tables. Περιλαμβάνεται επίσης το μετρήσιμο (measure) που εντοπίστηκε (αξία συναλλαγής - amount).

```
create table fact(  
  \pid int,  
  cardno char(16),  
  time_key datetime,  
  ttc int,  
  mcc int,  
  amount decimal(6,2),  
  primary key (pid, cardno, time_key, ttc, mcc),  
  foreign key (pid) references cardowner(pid),  
  foreign key (cardno) references cardinfo(cardno),  
  foreign key (time_key) references timeinfo(time_key),  
  foreign key (ttc) references transactiontype(ttc),  
  foreign key (mcc) references city(mcc),  
);
```

Οι απαιτήσεις της διοίκησης εστιάζουν στην
ανάλυση του αριθμού και της αξίας των
συναλλαγών...

↓
λαμβάνεται μέσω
της **count()**

↓
αποθηκεύεται ως amount

3. Τροφοδοτείται το σχήμα της αποθήκης με τα απαραίτητα στοιχεία από τον πίνακα CardsTransactions.

```
insert into cardinfo  
select distinct cardno, card_brand, card_type  
from CardsTransactions
```

```
insert into cardowner  
select distinct pid, pname, age, gender  
from CardsTransactions
```

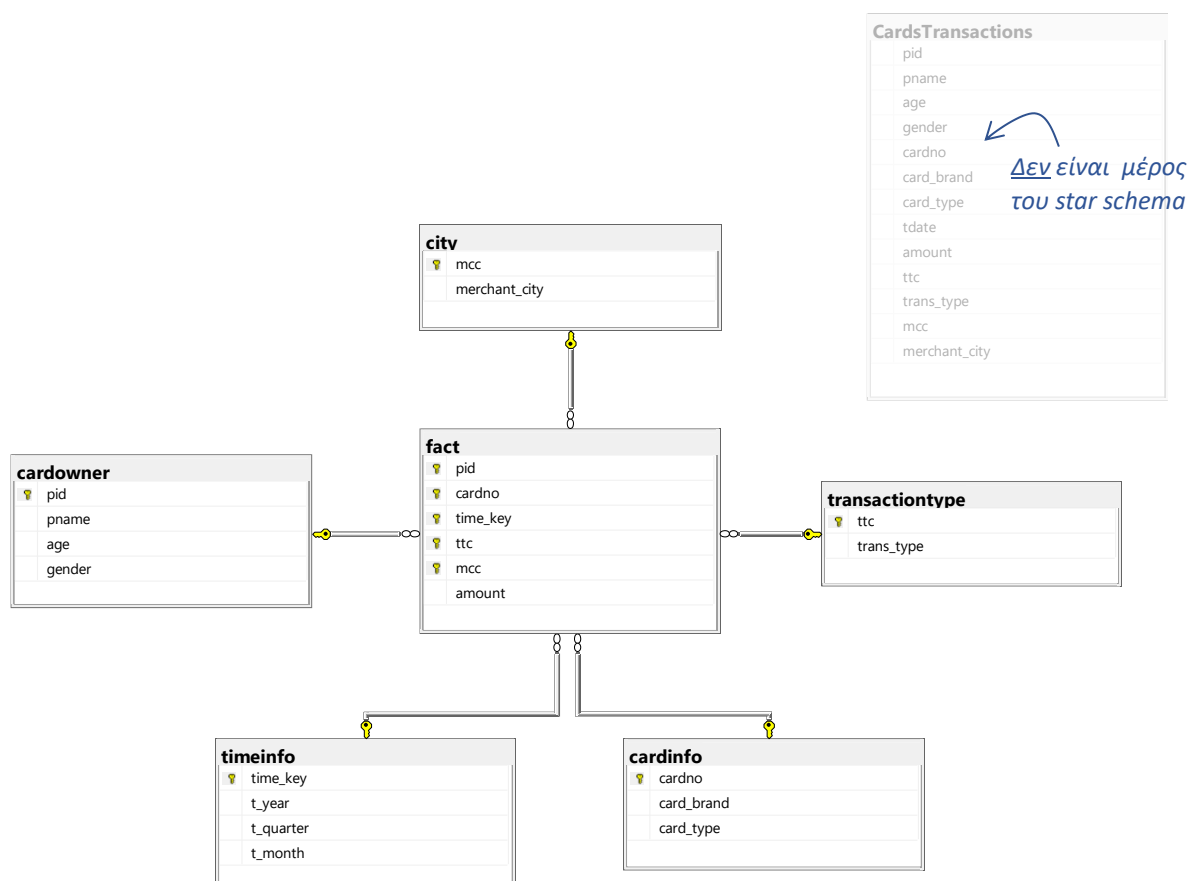
```
insert into transactiontype  
select distinct ttc, trans_type  
from CardsTransactions
```

```
insert into city  
select distinct mcc, merchant_city  
from CardsTransactions
```

```
insert into timeinfo  
select distinct tdate, datepart(year, tdate), datepart(quarter,  
tdate), datepart(month, tdate)  
from CardsTransactions
```

```
insert into fact  
select distinct pid, cardno, tdate, ttc, mcc, amount  
from CardsTransactions
```

4.



Ζήτηση 2^ο

1. Εμφανίστε έναν κατάλογο με την αξία των συναλλαγών ανά πόλη. Ο κατάλογος πρέπει να είναι ταξινομημένος με βάση την πόλη σε αύξουσα διάταξη.

```
select merchant_city as "Merchant City Code", sum(amount) as "Total Amount"
from fact, city
where fact.mcc = city.mcc
group by merchant_city
order by merchant_city
```

2. Εμφανίστε έναν κατάλογο με την αξία των συναλλαγών ανά έτος και φύλο. Ο κατάλογος πρέπει να είναι ταξινομημένος με βάση το έτος σε φθίνουσα διάταξη.

```
select t_year as "Year", gender as "Gender", sum(amount) as "Total Amount"
from fact, timeinfo, cardowner
where fact.time_key = timeinfo.time_key and fact.pid = cardowner.pid
group by t_year, gender
order by t_year desc
```

3. Εμφανίστε έναν κατάλογο με τον αριθμό και την αξία των συναλλαγών ανά επωνυμία (card_brand) είδος (card_type) κάρτας.

```
select card_brand as "Card brand", card_type as "Card type",  
count(amount) as "# of Transactions", sum(amount) as "Total amount"  
from fact, cardinfo  
where fact.cardno = cardinfo.cardno  
group by card_brand, card_type
```

4. Εμφανίστε έναν κατάλογο με ανάλυση της αξίας των συναλλαγών ανά τύπο συναλλαγής (trans_type) σε τριμηνιαία βάση για το έτος 2019.

Σημείωση: Επειδή ζητείται ανάλυση της αξίας των συναλλαγών, επιλέχθηκε ο τελεστής rollup για να παρουσιαστούν οι εξής πληροφορίες:

- α) συνολική αξία των συναλλαγών του έτους 2019.
- β) αξία ανά τύπο συναλλαγής για το έτος 2019.
- γ) αξία ανά τύπο συναλλαγής και τρίμηνο για το έτος 2019.

```
select trans_type as "Transaction type", t_quarter as "Quarter",  
sum(amount) as "Total amount"  
from fact, transactiontype, timeinfo  
where fact.ttc = transactiontype.ttc and  
fact.time_key = timeinfo.time_key and t_year = 2019  
group by rollup(trans_type, t_quarter)
```

5. Η διοίκηση της τράπεζας θέλει μία αναφορά που θα περιέχει τις ακόλουθες πληροφορίες για τις online συναλλαγές (trans_type='Online Transaction').

- a. Την συνολική αξία των online συναλλαγών.
- b. Την αξία των online συναλλαγών ανά έτος.
- c. Την αξία των online συναλλαγών ανά έτος και φύλο.
- d. Την αξία των online συναλλαγών ανά έτος, φύλο και ηλικία.

```
select t_year as "Year", gender as "Gender", age as "Age",  
sum(amount) as "Total Amount"  
from fact, transactiontype, timeinfo, cardowner  
where fact.ttc = transactiontype.ttc and  
trans_type = 'Online Transaction' and  
fact.time_key = timeinfo.time_key and fact.pid = cardowner.pid  
group by rollup(t_year, gender, age)
```

Ζήτημα 3^ο

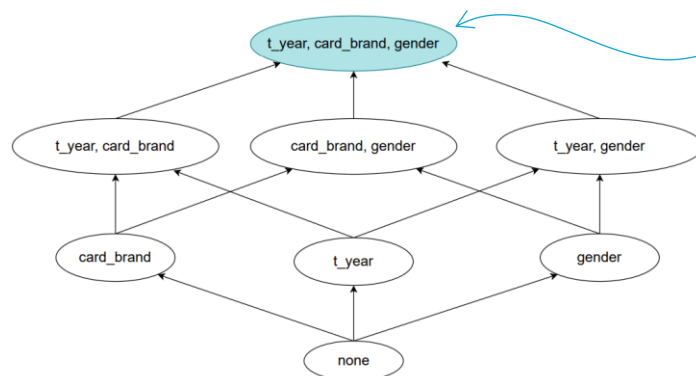
1. Γράψτε μια επερώτηση σε γλώσσα SQL το αποτέλεσμα της οποίας είναι η δημιουργία ενός κύβου (data cube), κάθε κελί του οποίου περιέχει τον αριθμό των συναλλαγών για έναν συγκεκριμένο συνδυασμό τιμών: έτος, επωνυμία κάρτας (card_brand) και φύλο.

```
select t_year as "Year", card_brand as "Card Brand", gender as
"Gender", count(amount) as "# of Transactions"
from fact, timeinfo, cardinfo, cardowner
where fact.time_key = timeinfo.time_key and fact.cardno =
cardinfo.cardno and fact.pid = cardowner.pid
group by cube(t_year, card_brand, cardowner.gender)
```

2. Ζητείται η δημιουργία μιας MATERIALIZED όψης η οποία θα περιέχει το αποτέλεσμα ενός μόνο GROUP BY του κύβου του ερωτήματος 1 και θα αρκεί για την παραγωγή των υπόλοιπων BY του κύβου από την ίδια την όψη αυτή.

Εφόσον ο κύβος θα έχει 3 διαστάσεις και χρησιμοποιείται 1 γνώρισμα για κάθε διάσταση, τα δυνατά GROUP BYs που προκύπτουν είναι $2^3 = 8$ το πλήθος.

Το 3D Data Cube Lattice που προκύπτει για τον κύβο του προηγούμενου ερωτήματος και αναδεικνύει τις εξαρτήσεις μεταξύ των δυνατών GROUP BYs είναι το εξής:



Όπως είναι αναμενόμενο και γίνεται αντιληπτό πιο παραστατικά από το 3D Data Cube Lattice, υπάρχει μόνο ένα GROUP BY από το οποίο δύναται να υπολογιστούν τα υπόλοιπα και αυτό είναι το πιο λεπτομερές: **t_year, card_brand, gender**

Συνεπώς, η όψη θα περιέχει το αποτέλεσμα του GROUP BY **t_year, card_brand, gender**.

Παρακάτω δημιουργείται η MATERIALIZED όψη **v_year_cardbrand_gender**:

```
create view v_year_cardbrand_gender
with schemabinding as
select t_year, card_brand, gender, count_big(*) as
count_transactions
from dbo.fact, dbo.timeinfo, dbo.cardinfo, dbo.cardowner
where fact.time_key = timeinfo.time_key and fact.cardno =
cardinfo.cardno and fact.pid = cardowner.pid
group by t_year, card_brand, gender
```

```
create unique clustered index uni_idx_v_year_cardbrand_gender
on v_year_cardbrand_gender(t_year, card_brand, gender)
```

Μετά από τη δημιουργία της όψης θα πρέπει να εκτελεστούν οι εντολές για 7 GROUP BYs (καθώς για το GROUP BY που αντιστοιχεί στην όψη που δημιουργήθηκε αρκεί να γίνει select):

t_year, card_brand, gender

```
select t_year as "Year", card_brand as "Card Brand", gender as  
"Gender", count_transactions as "# of Transactions"  
from v_year_cardbrand_gender
```

t_year, card_brand

```
select t_year as "Year", card_brand as "Card Brand",  
sum(count_transactions) as "# of Transactions"  
from v_year_cardbrand_gender  
group by t_year, card_brand
```

t_year, gender

```
select t_year as "Year", gender as "Gender",  
sum(count_transactions) as "# of Transactions"  
from v_year_cardbrand_gender  
group by t_year, gender
```

card_brand, gender

```
select card_brand as "Card Brand", gender as "Gender",  
sum(count_transactions) as "# of Transactions"  
from v_year_cardbrand_gender  
group by card_brand, gender
```

t_year

```
select t_year as "Year", sum(count_transactions) as "# of  
Transactions"  
from v_year_cardbrand_gender  
group by t_year
```

card_brand

```
select card_brand as "Card Brand", sum(count_transactions) as  
"# of Transactions"  
from v_year_cardbrand_gender  
group by card_brand
```

gender

```
select gender as "Gender", sum(count_transactions) as "# of  
Transactions"  
from v_year_cardbrand_gender  
group by gender
```

none

```
select sum(count_transactions) as "Total # of Transactions"  
from v_year_cardbrand_gender
```

CTDW Bank

Συναλλαγές 2015-2020

1

Εφαρμόστε φίλτρα.

Έτος

2015

2020

Φύλο

☐ F

☐ M

Επωνυμία κάρτας

☐ Amex

☐ Discover

☐ Mastercard

☐ Visa

Κύβος πλήθους συναλλαγών

t_year	F	M	Σύνολο
2020	87816	79901	167717
Visa	32817	29433	62250
Mastercard	48346	43287	91633
Discover	2858	1768	4626
Amex	3795	5413	9208
2019	452398	397572	849970
Visa	165232	152154	317386
Mastercard	248422	207863	456285
Discover	17646	8899	26545
Amex	21098	28656	49754
2018	451111	395869	846980
Visa	165538	149293	314831
Mastercard	246959	210773	457732
Discover	18155	7931	26086
Amex	20459	27872	48331
2017	447347	396414	843761
Visa	164034	151196	315230
Mastercard	246099	210558	456657
Discover	17785	7584	25369
Amex	19429	27076	46505
Σύνολο	2331513	2055346	4386859

4,39 εκατ.

Πλήθος amount

235,70 εκατ.

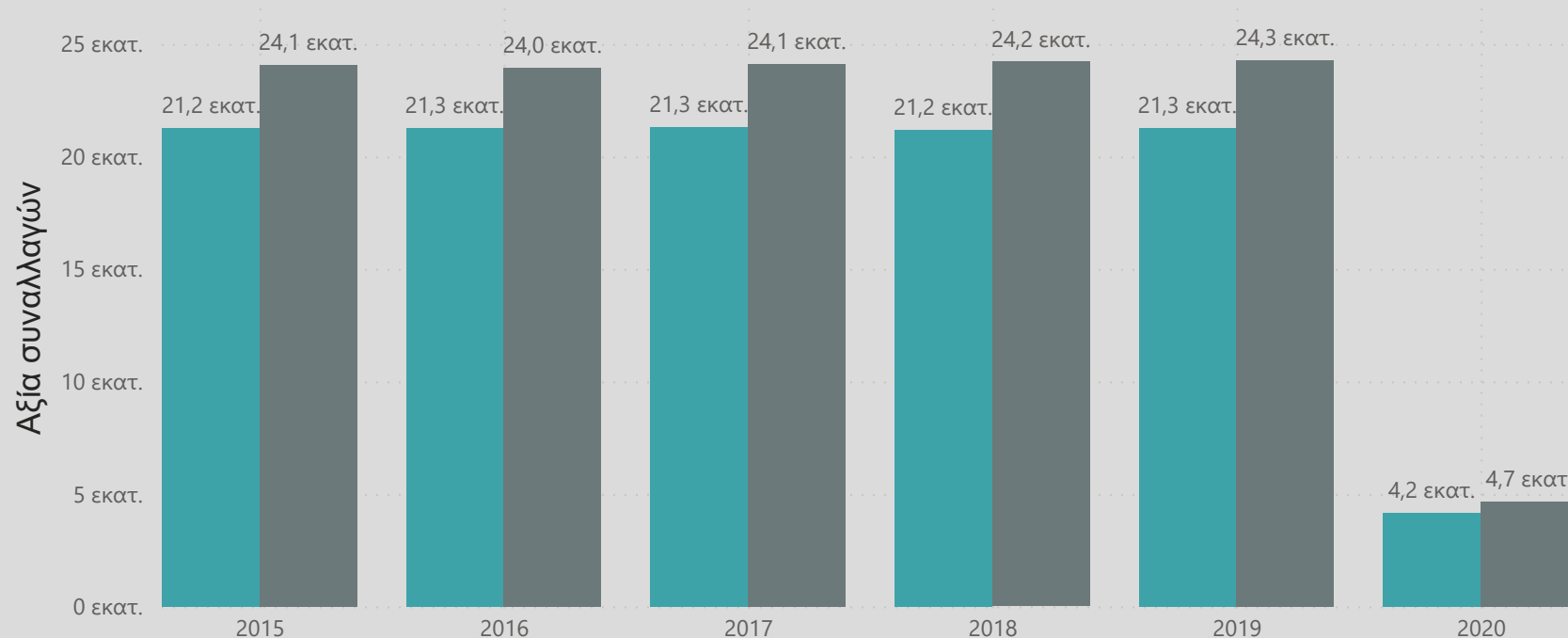
Άθροισμα amount

2

Περιηγηθείτε στα αποτελέσματα.

Συνολική αξία των συναλλαγών ανά έτος και φύλο

φύλο ● M ● F



CTDW Bank

Συναλλαγές 2015-2020

Έτος

2015 2020



Φύλο

F

M

Επωνυμία κάρτας

Amex

Discover

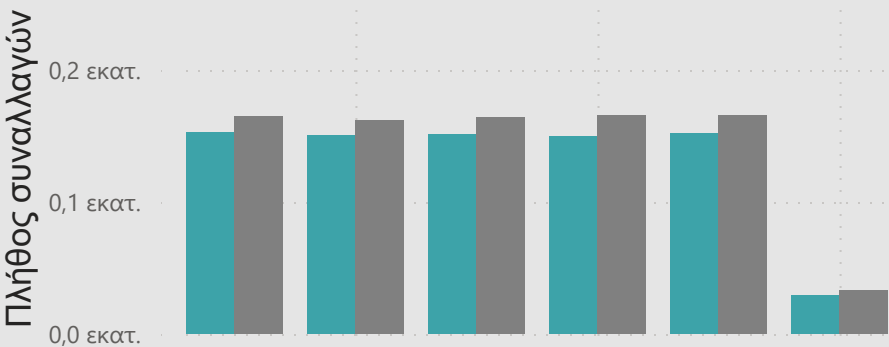
Mastercard

Visa

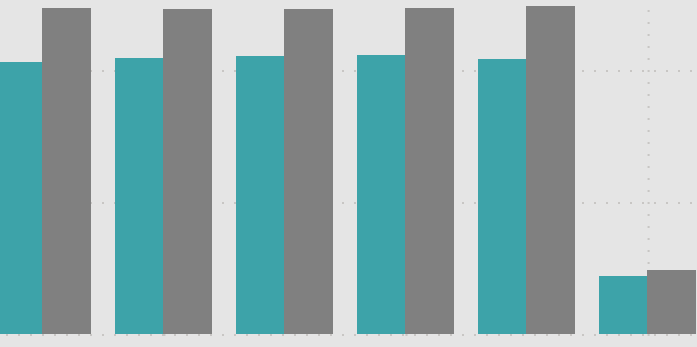
Κύβος πλήθους συναλλαγών

φύλο M F

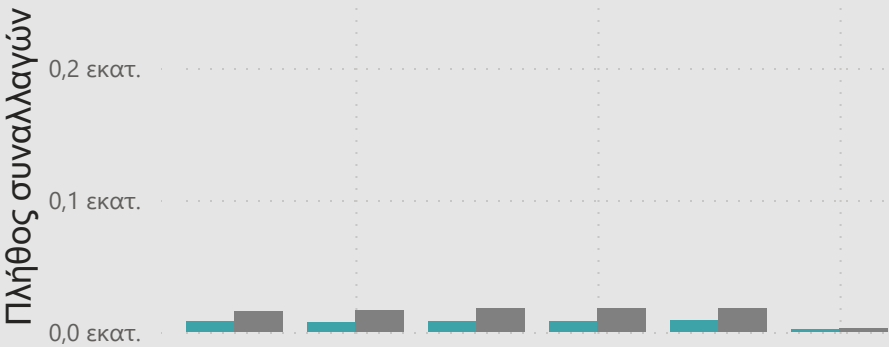
Visa



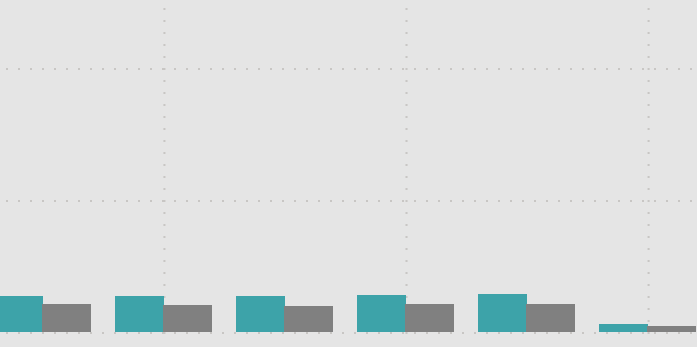
Mastercard



Discover



Amex



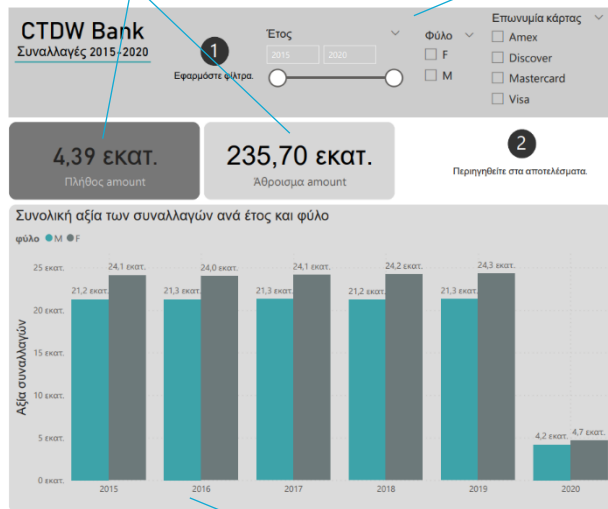
Κύβος πλήθους συναλλαγών

card_brand	F	M	Σύνολο
Visa	854857	784503	1639360
2020	32817	29433	62250
2019	165232	152154	317386
2018	165538	149293	314831
2017	164034	151196	315230
2016	162114	150106	312220
2015	165122	152321	317443
Mastercard	1282850	1086617	2369467
2020	48346	43287	91633
2019	248422	207863	456285
2018	246959	210773	457732
2017	246099	210558	456657
2016	246079	208798	454877
2015	246945	205338	452283
Discover	88316	41389	129705
2020	2858	1768	4626
2019	17646	8899	26545
2018	18155	7931	26086
2017	17785	7584	25369
2016	16119	7487	23606
Σύνολο	2331513	2055346	4386859

Παρατηρήσεις:

Πλήθος και αξία συναλλαγών για τα επιλεγμένα φίλτρα (συνολικό πλήθος και αξία αν δεν έχουν εφαρμοστεί φίλτρα).

Φίλτρα για τις διαστάσεις που εμφανίζονται στα δεδομένα που απεικονίζονται: έτος, φύλο, τύπος κάρτας.



Κύβος πλήθους συναλλαγών			
t_year	F	M	Σύνολο
2020	87816	79901	167717
Visa	32817	29433	62250
Mastercard	48346	43287	91633
Discover	2858	1768	4626
Amex	3795	5413	9208
2019	452398	397572	849970
Visa	165232	152154	317386
Mastercard	248422	207863	456285
Discover	17646	8899	26545
Amex	21098	28656	49754
2018	451111	395869	846980
Visa	165538	149293	314831
Mastercard	246959	210773	457732
Discover	18155	7931	26086
Amex	20459	27872	48331
2017	447347	396414	843761
Visa	164034	151196	315230
Mastercard	246099	210558	456657
Discover	17785	7584	25369
Amex	19429	27076	46505
Σύνολο	2331513	2055346	4386859

2^ο ζητούμενο – πίνακας/μήτρα:

Απεικόνιση του κύβου (data cube) του πλήθους των συναλλαγών ανά έτος, τύπο κάρτας και φύλο.

Γραμμές

t_year

card_brand

Στήλες

gender

Τιμές

Πλήθος amount

Άξονας x

t_year

Άξονας y

Αθροισμα amount

Υπόμνημα

gender

1^ο ζητούμενο - ραβδόγραμμα: Απεικόνιση της αξίας των συναλλαγών ανά έτος και φύλο.



Κύβος πλήθους συναλλαγών			
card_brand	F	M	Σύνολο
2020	854857	784503	1639360
Visa	32817	29433	62250
Mastercard	48346	43287	91633
Discover	2858	1768	4626
Amex	3795	5413	9208
2019	452398	397572	849970
Visa	165232	152154	317386
Mastercard	248422	207863	456285
Discover	17646	8899	26545
Amex	21098	28656	49754
2018	451111	395869	846980
Visa	165538	149293	314831
Mastercard	246959	210773	457732
Discover	18155	7931	26086
Amex	20459	27872	48331
2017	447347	396414	843761
Visa	164034	151196	315230
Mastercard	246099	210558	456657
Discover	17785	7584	25369
Amex	19429	27076	46505
Σύνολο	2331513	2055346	4386859

2^ο ζητούμενο – πίνακας/μήτρα:

Απεικόνιση του κύβου (data cube) του πλήθους των συναλλαγών ανά τύπο κάρτας, έτος και φύλο.

Γραμμές

card_brand

t_year

Στήλες

gender

Τιμές

Πλήθος amount

Άξονας x

t_year

Άξονας y

Πλήθος amount

Υπόμνημα

gender

Μικρά πολλαπλά

card_brand

2^ο ζητούμενο – ραβδογράμματα:

Απεικόνιση του κύβου (data cube) του πλήθους των συναλλαγών ανά τύπο κάρτας, έτος και φύλο.