

# Bank transactions analysis

# Plan

- **1- Introduction**
- **2- Problématique et objectifs**
- **3- Données**
- **4- Technologies utilisées**
- **5- Data Intégration ETL avec Talend**
- **6- Data Visualisation et Analyses**
- **7- DashBoard**
- **8- Conclusion**

# 1- Introduction

- **Petr Berka a recensé les données d'une Banque de la republique Tchèque**
- **Données des clients et leurs transactions**
- **5369 échantillon de clients ont été recensé sur la période de 1993 à 1998**

## 2- Problématique et objectifs

- **Comprendre le comportement client sur ses transactions en faisant des analytiques**
- **Construire un dashboard de visualisation contenant des graphs et KPI**

## 3- Données

- **Donnée accessible sur Data World**
- **Composée de huit tables**
  - Client.csv
  - Account.csv
  - Card.csv
  - Transaction.csv
  - Loan.csv, District.csv
  - Order.csv, Disposition.csv,

# Table Client.csv



client.csv

Client

View



	#	client_id	▼	#	birth_number	#	district_id	▼	
1		1			706213		18		
2		2			450204		1		
3		3			406009		1		
4		4			561201		5		
5		5			605703		5		

Showing 1-5 of 5,369 rows, 3 columns [See all](#)

[Switch to column overview](#)

# Table Transaction.csv



trans.csv  
Transaction

View



	#	trans_id	▼	#	account_id	▼	#	date	T	type	T	operation	#	amount	#	balance
1		695247			2378			930101		PRIJEM		VKLAD		700.00		700.00
2		171812			576			930101		PRIJEM		VKLAD		900.00		900.00
3		207264			704			930101		PRIJEM		VKLAD		1000.00		1000.00
4		1117247			3818			930101		PRIJEM		VKLAD		600.00		600.00
5		579373			1972			930102		PRIJEM		VKLAD		400.00		400.00

Showing 1-5 of 1,056,320 rows, 10 columns See all

Switch to column overview

## 4- Technologies utilisées

- **Talend**
- **Tableau Desktop**



# 5- Data Intégration ETL avec Talend

- **Data processing**

- Extraire la date de naissance et le sexe à partir d'une colonne birth\_number de la table Client.csv
- Dans cette table birth\_number est au type Int et sous la forme YYMMDD pour les hommes et YYMM(50+DD) pour les femmes
- Pour les autres tables qui contiennent des dates elles sont au type Int (YYMMDD)

- **Renommer les colonnes**

# Talend project

Talend Open Studio for Data Integration (7.3.1.20200219\_1130) | Bank\_Transaction (Connection: Local)

File Edit View Window Help

Learn Ask Exchange Videos Cloud

Repository

LOCAL: Bank\_Transaction

- Business Models
- Job Designs
  - Bank\_Data\_Processing 0.1
- Contexts
- Code
- SQL Templates
- Metadata
  - Db Connections
  - File delimited
    - Account 0.1
    - Card 0.1
    - Clients 0.1
    - Disp 0.1
    - District 0.1
    - Loan 0.1
    - Order 0.1
    - Transaction 0.1
  - File positional
  - File regex

Outline Code View...

- tFileInputDelimited\_1 (Disposition)
- tFileInputDelimited\_2 (Clients)
- tFileInputDelimited\_3 (Account)
- tFileInputDelimited\_4 (District)
- tFileInputDelimited\_5 (Card)

Job Bank\_Data\_Processing 0.1

Designer Code

Job(Bank\_Data\_Processing 0.1) Contexts(Bank\_Data\_Processing) Component Run (Job Bank\_Data\_Processing)

SubJob

Basic settings

Restore default color Save as default color

**Clients**  
5369 rows in 0.63s  
5369 rows in 0.63s  
563 rows/s  
row1 (Main)  
process1 (Main)  
tMap\_1  
Client\_output

**Disposition**  
5369 rows in 0.11s  
47513.27 rows/s  
row2 (Main)  
Disposition  
Disposition\_output

**Account**  
4500 rows in 0.19s  
4500 rows in 0.19s  
3684.21 rows/s  
row3 (Main)  
process2 (Main)  
tMap\_2  
Account\_output

**District**  
76 rows in 0.07s  
76 rows in 0.07s  
1134.33 rows/s  
row4 (Main)  
out1 (Main)  
tMap\_3  
District\_output

**Card**  
892 rows in 0.07s  
892 rows in 0.07s  
12388.89 rows/s  
row5 (Main)  
out2 (Main)  
tMap\_4  
Card\_output

**Order**  
6471 rows in 0.16s  
40192.55 rows/s  
row6 (Main)  
Order  
Order\_output

**Transaction**  
1056320 rows in 8.8s  
1056320 rows in 8.8s  
10036.26 rows/s  
row7 (Main)  
out3 (Main)  
tMap\_5  
Transaction\_output

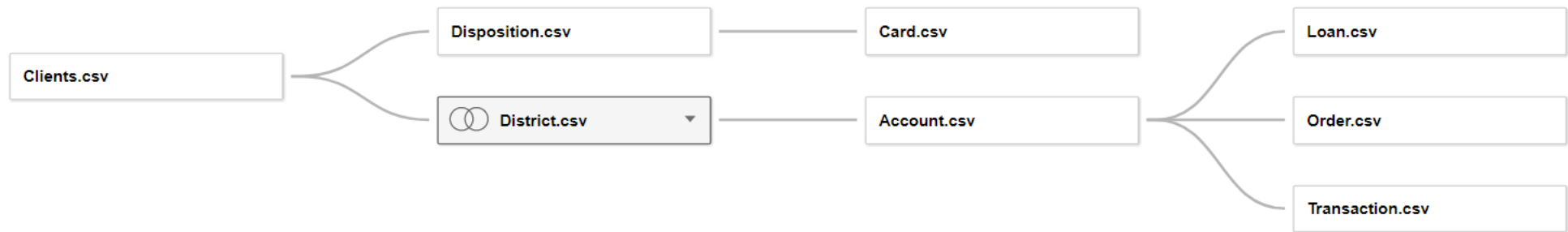
**Loan**  
682 rows in 0.01s  
682 rows in 0.01s  
85250 rows/s  
row8 (Main)  
out4 (Main)  
tMap\_6  
Loan\_output

# Donnée sur Tableau desktop

Clients.csv+ (Multiple Connections)

Connection  
☒ Live ☐ Extract

Filters  
 0 | [Add](#)



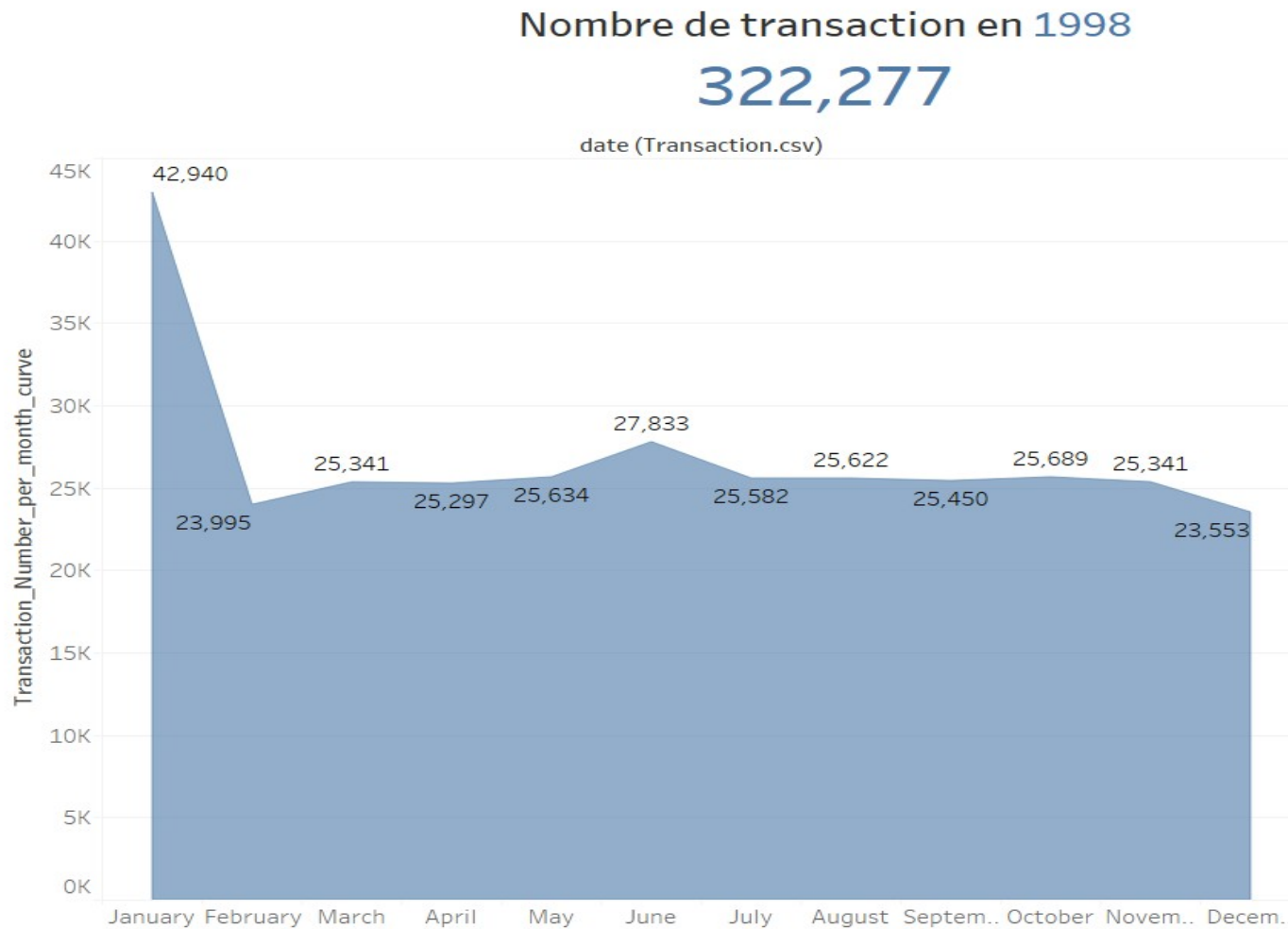
Sort fields: Data source order

☐ Show aliases ☐ Show hidden fields 1,000 rows

#	Clients.csv1	Abc	#	#	Abc	Abc	#	#	#	#
	client id (Clien...	birthday (Clie...	gender (Client...	district id (Cli...	District Code	District Name	Region	Nb Inhabitant	Nb Municipa L...	Nb Municipa B...
	1	12/13/70	Female	18	18	Pisek	south Bohemia	70,699	60	13
	2	02/04/45	Male	1	1	Hl.m. Praha	Prague	1,204,953	0	0
	3	10/09/40	Female	1	1	Hl.m. Praha	Prague	1,204,953	0	0
	4	12/01/56	Male	5	5	Kolin	central Bohemia	95,616	65	30
	5	07/03/60	Female	5	5	Kolin	central Bohemia	95,616	65	30
	6	09/22/19	Male	12	12	Pribram	central Bohemia	107,870	84	29
	7	01/25/29	Male	15	15	Cesky Krumlov	south Bohemia	58,796	22	16
	8	02/21/38	Female	51	51	Trutnov	east Bohemia	121,947	37	28
	9	10/16/35	Male	60	60	Prostejov	south Moravia	110,643	49	41
	10	05/01/43	Male	57	57	Hodonin	south Moravia	161,954	21	37

## 6- Data Visualisation

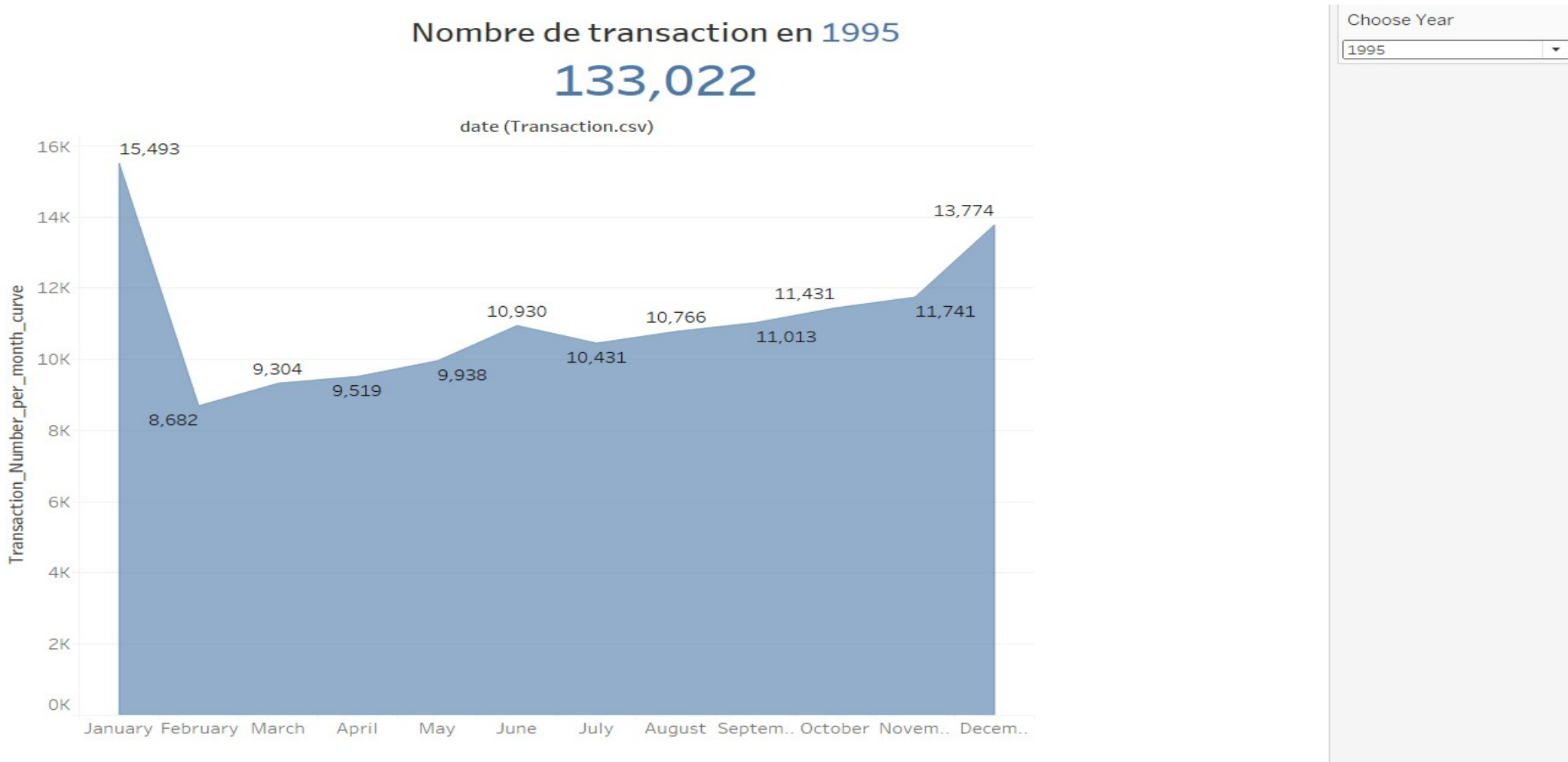
# Nombre de transactions effectuées en fonction du mois (1)



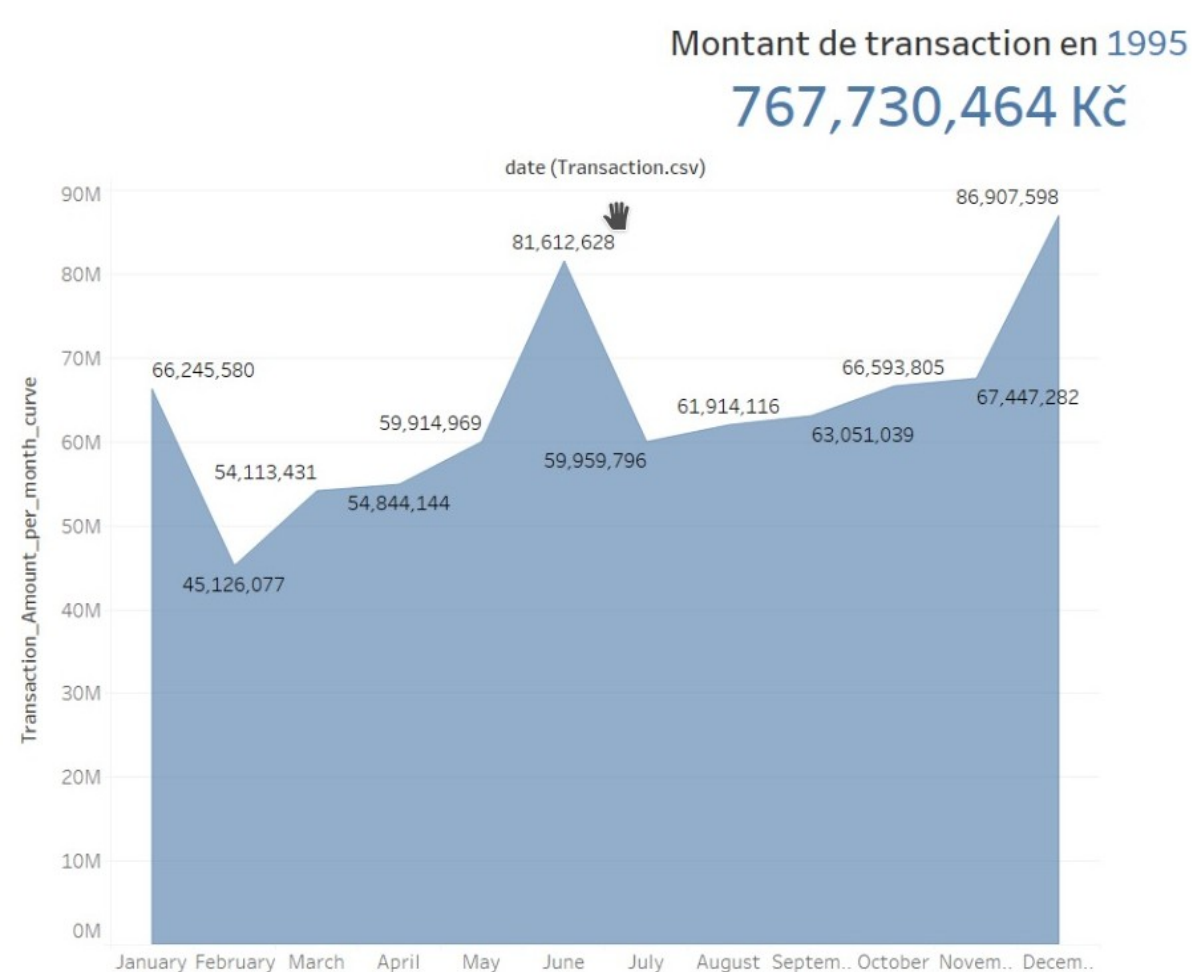
Choose Year

1998

# Nombre de transactions effectuées en fonction du mois (2)



# Montant de transaction en fonction de l'année en monnaie Tchèque ( Koruna: Kč)



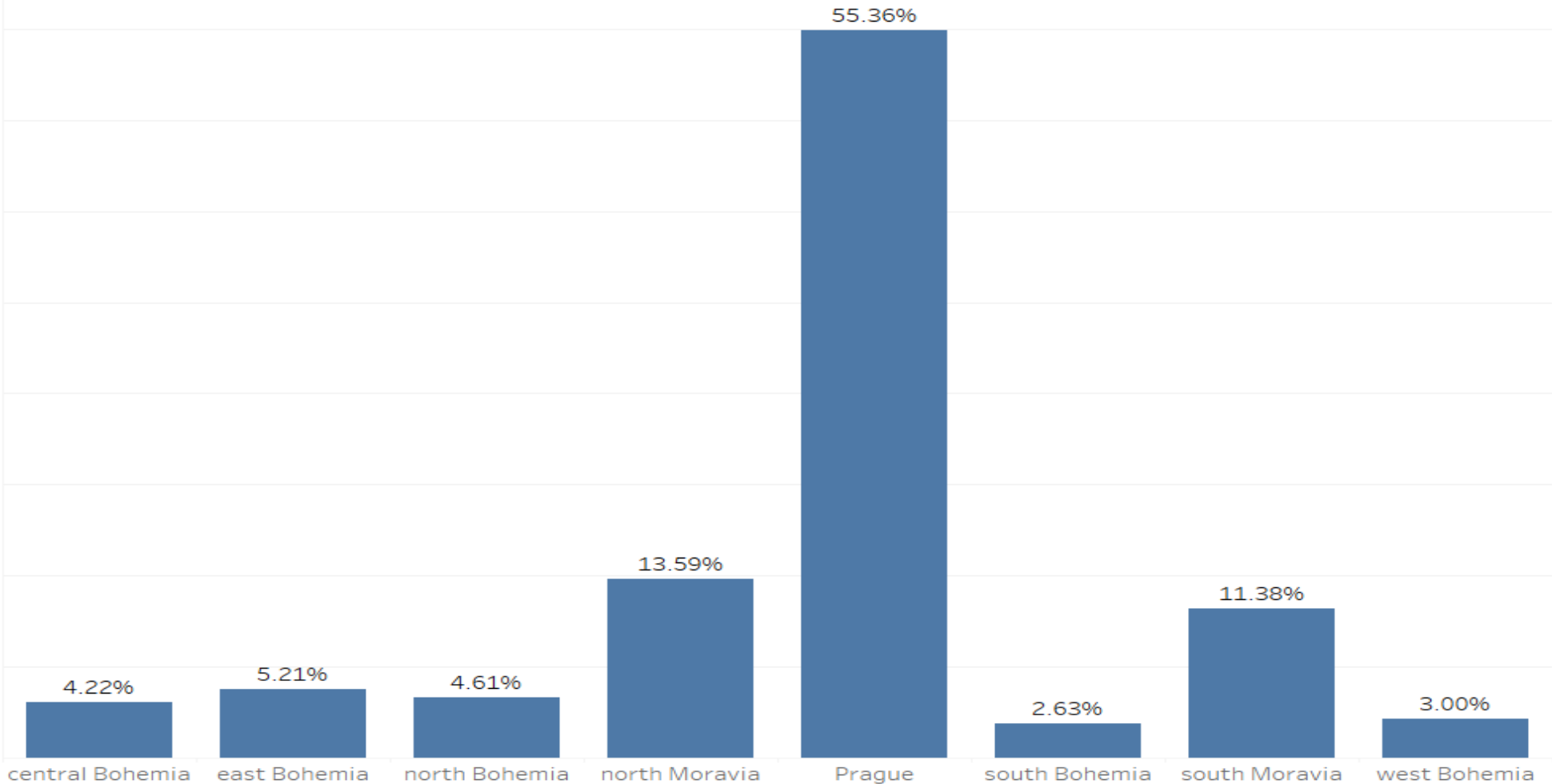
Choose Year

1995

# Part de population dans chaque région du pays

Nombre d'habitant par régions

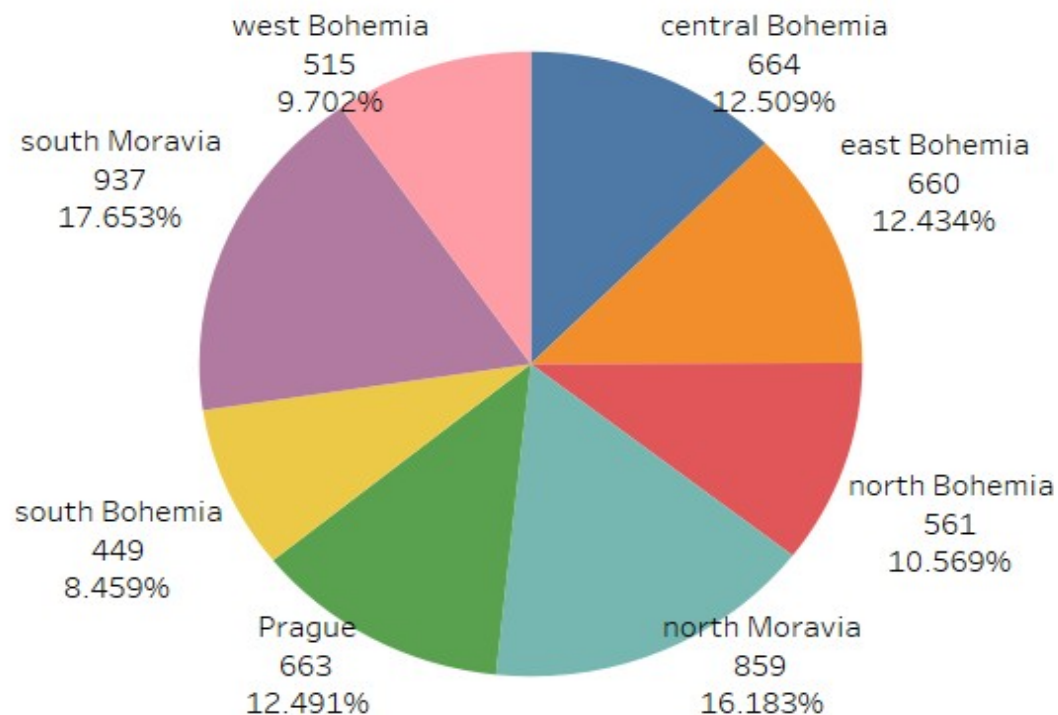
Region





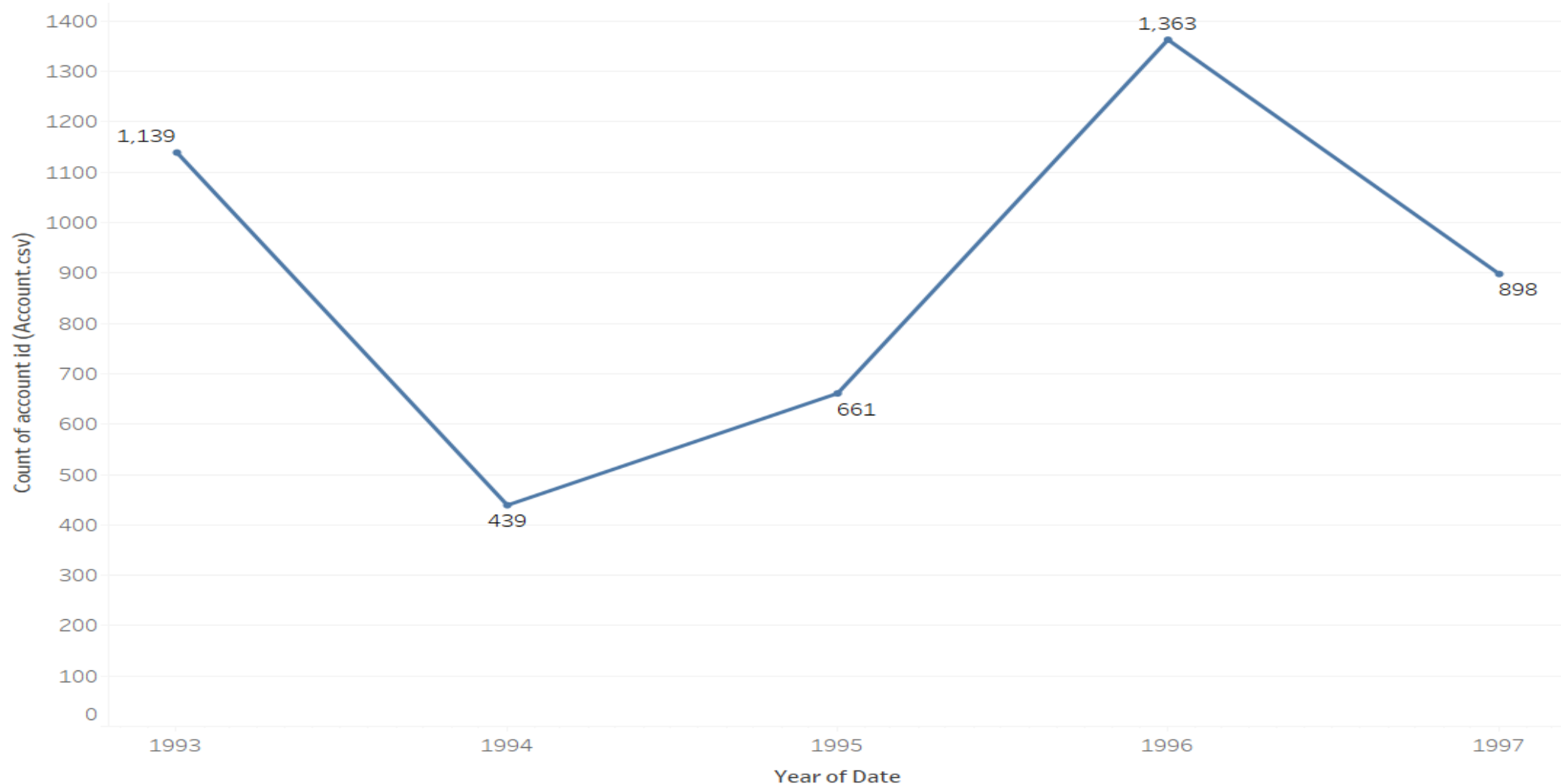
# Part des Clients de la Banque en fonction de la région

Part des Clients par Région

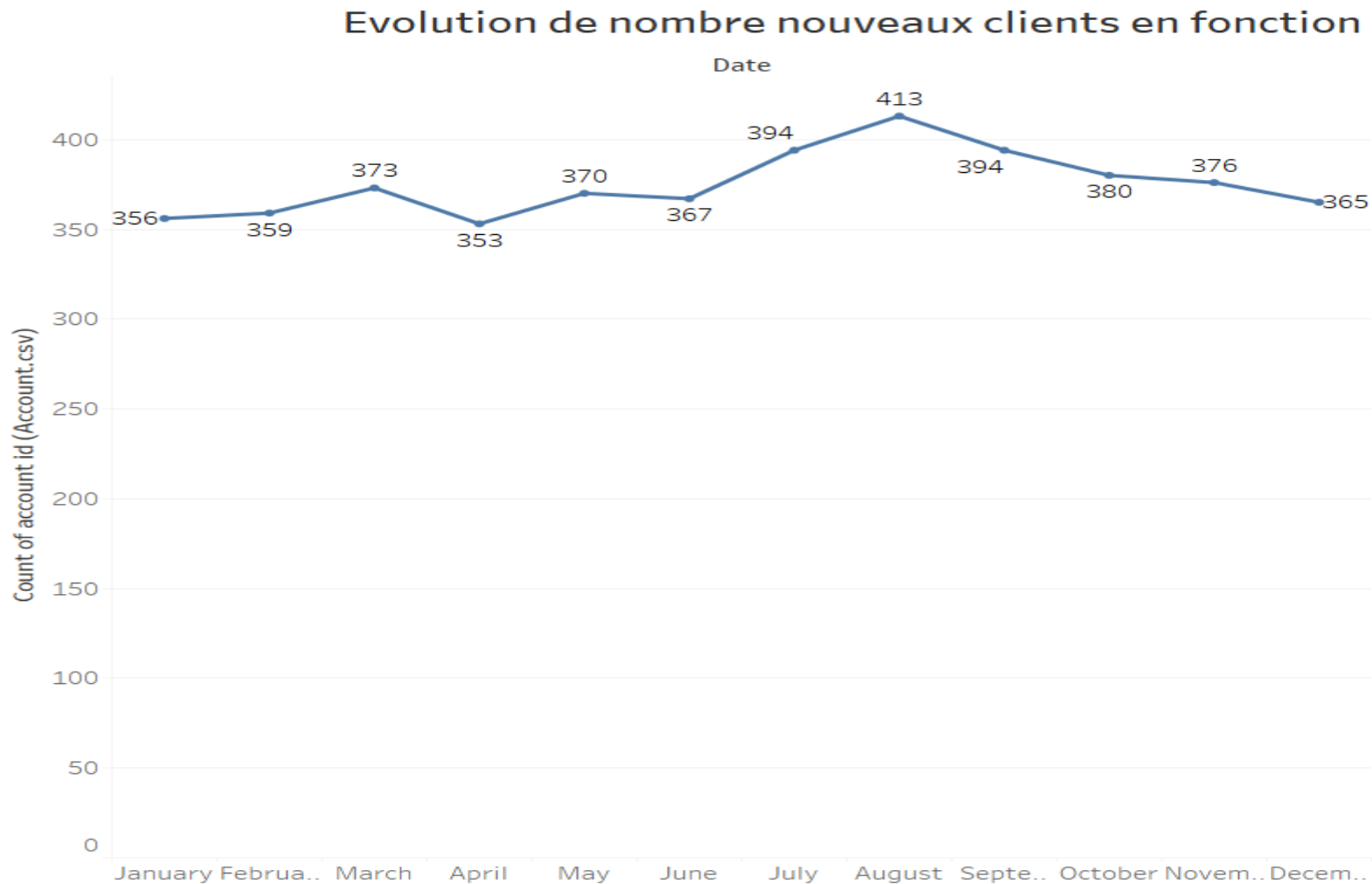


# Evolution de nouveaux clients en fonction de l'année

Evolution de nombre nouveaux clients en fonction des années



# Evolution de nouveaux clients en fonction du mois



# DashBoard



## Tchèque Bank DashBoard

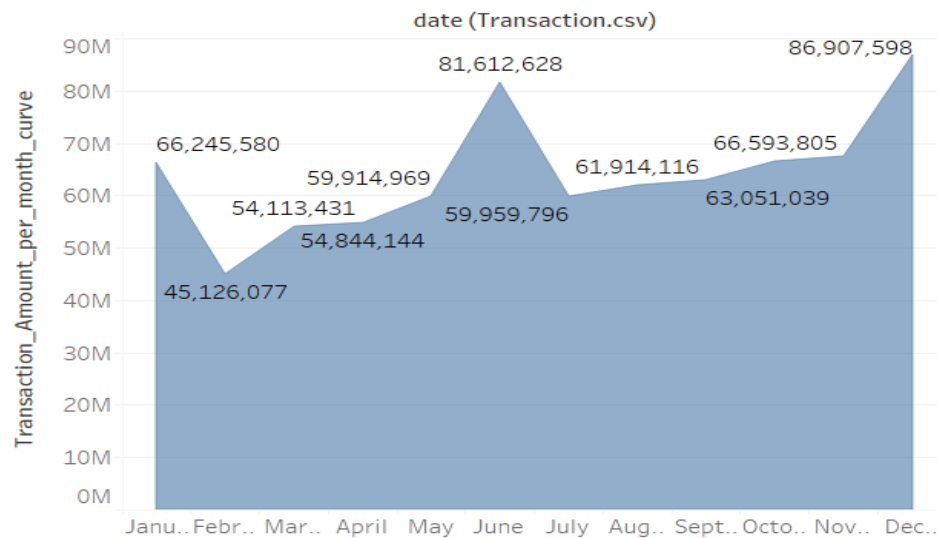
Choose Year  
1995

Region  
central Bohemian  
east Bohemian  
north Bohemian  
north Moravian  
Prague  
south Bohemian  
south Moravian  
west Bohemian

Evolution New ...  
Mois

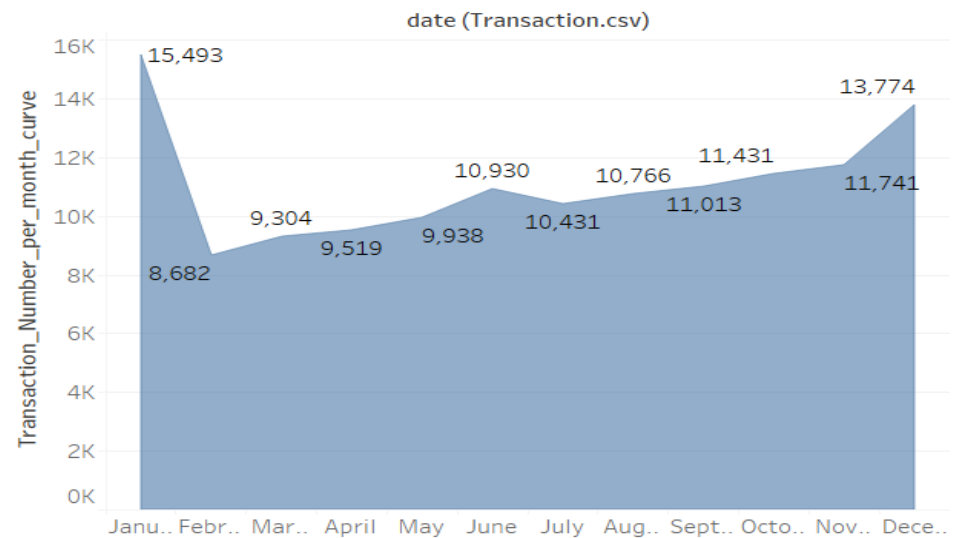
Montant de transaction en 1995

767,730,464 Kč



Nombre de transaction en 1995

133,022



Part des Clients par Région

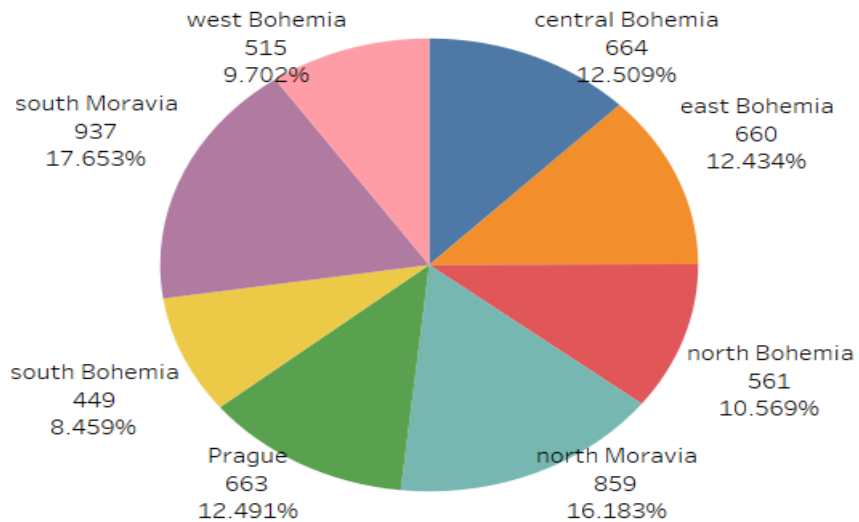
Nombre d'habitant par régions

# DashBoard

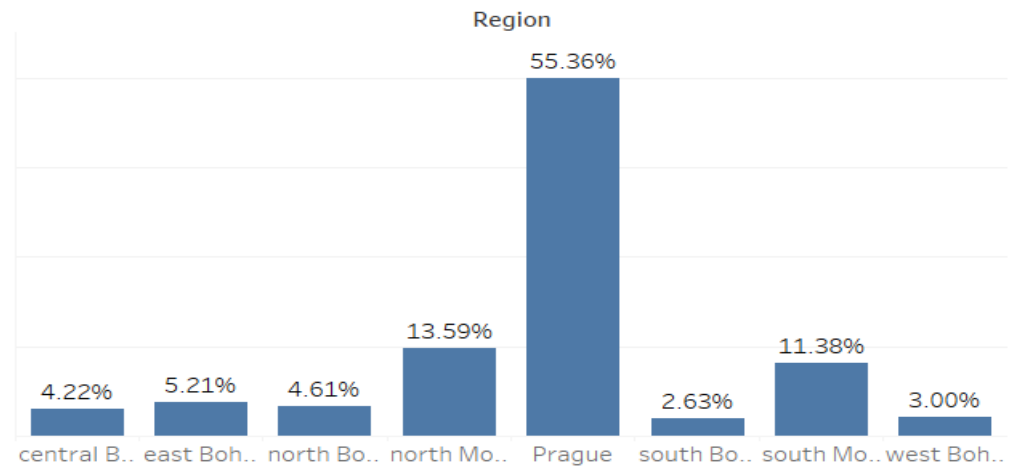
Janv. Févr. Mars. Avril. Mai. Juin. July. Août. Sept. Octo. Nov. Dec.

Janv. Févr. Mars. Avril. Mai. Juin. July. Août. Sept. Octo. Nov. Dec.

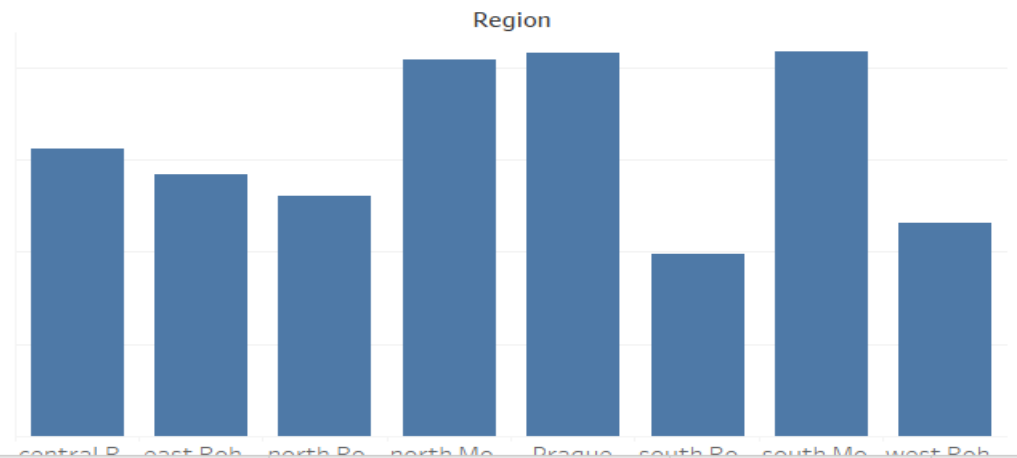
Part des Clients par Région



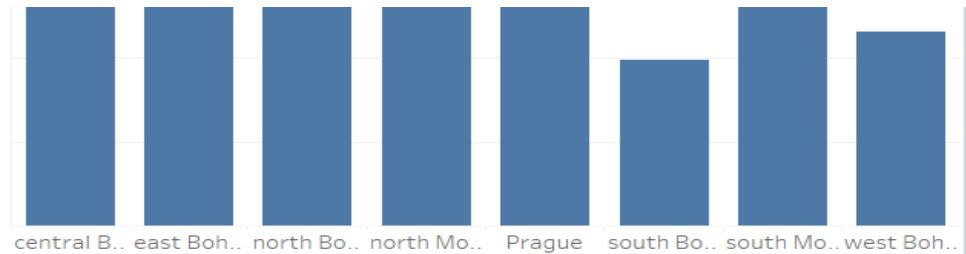
Nombre d'habitant par régions



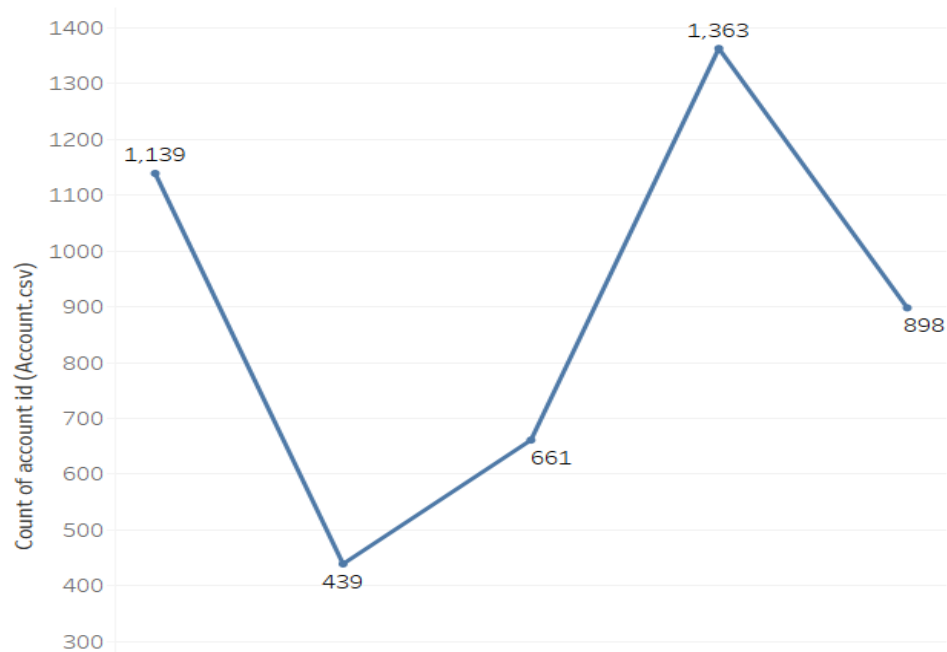
Moyenne des salaires en fonction des régions



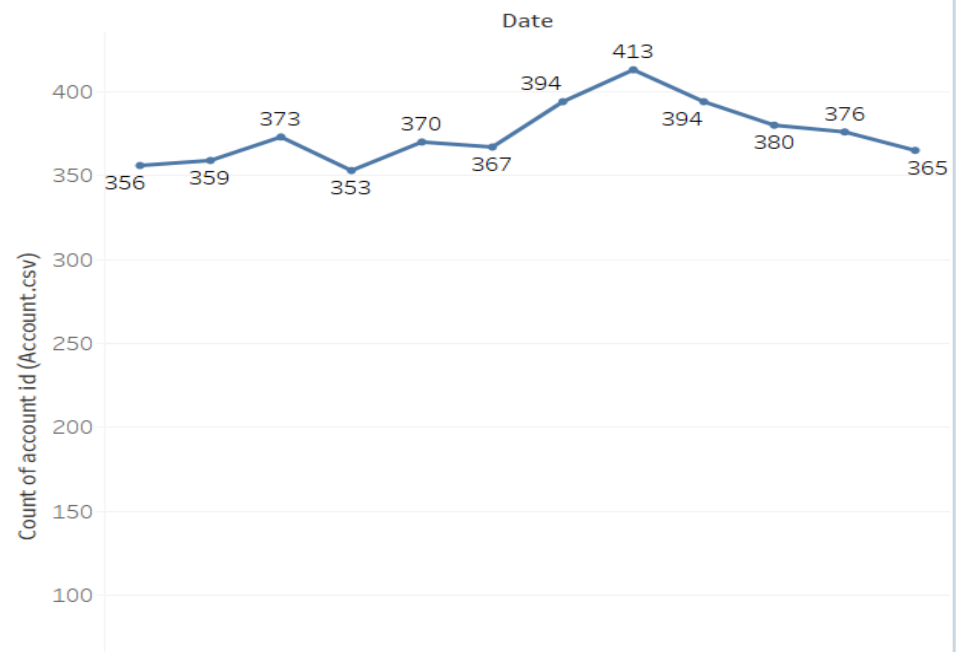
# DashBoard



Evolution de nombre nouveaux clients en fonction des années



Evolution de nombre nouveaux clients en fonction du mois



# Conclusion

**Merci pour votre  
attention...**