

# **Business Intelligence**

**Analiza magazinelor de patiserie și  
produse de panificație**

**Proiect realizat de**

Stratu Mihaela

# Cuprins

<b>Problema de Business.....</b>	<b>3</b>
<b>Metodologie de Lucru .....</b>	<b>4</b>
<b>Modelarea setului de date.....</b>	<b>5</b>
<b>Vizualizarea și Analiza Datelor .....</b>	<b>9</b>
<b>Topul firmelor .....</b>	<b>9</b>
<b>Comparații .....</b>	<b>10</b>
<b>Evoluția numărului de firme în funcție de mediu.....</b>	<b>11</b>
<b>Procent Cheltuieli din Venituri .....</b>	<b>12</b>
<b>Analiza de trend pentru numărul de salariați .....</b>	<b>13</b>
<b>Harta.....</b>	<b>14</b>
<b>Dashboard .....</b>	<b>15</b>
<b>Story .....</b>	<b>16</b>

# Problema de Business

Proiectul de față își propune să realizeze o analiză de piață a magazinelor din județul Cluj ce au ca și specific vânzarea de produse de patiserie și panificație. Perioada luată în calcul pentru acest studiu este 2018/2022. Tematica acestei lucrări a fost aleasă din pasiunea pentru dulciuri a fiecăruia dintre noi, sedimentată de expansiunea acestui tip de magazine în ultima perioadă.

În vederea acestui proiect am luat ca și referință codurile CAEN 1071, 1072 și 1073, utilizate ca și identificatori specifici pentru fabricarea produselor de brutărie și produselor făinoase, codul de bază pentru aceste categorii fiind 107X.

Platformele utilizate în contextul acestei analize au fost Tableau Prep pentru prelucrarea de date și Tableau Desktop pentru vizualizarea grafică și structurată a datelor. Prin îmbinarea celor două tehnologii am reușit să creăm un studiu de caz orientat pe satisfacerea necesităților utilizatorilor.



# Metodologie de Lucru

Pentru realizarea acestui proiect am colaborat ca o echipă stabilind obiectivele de analiză în cadrul primei întâlniri, urmând a distribui ulterior sarcinile în mod egal.

În urma alegerii temei am început lucrul propriu zis pentru selectarea setului de date, luarea de decizii de modelare, documentând în paralel fiecare pas al acestui proces. La final am realizat vizualizările și am revizuit informațiile prelucrate, alături de documentul final.

Colaborarea a fost una de excepție și am sedimentat aptitudinile de analiză și interpretare de date învățate de-a lungul timpului.

# Modelarea setului de date

Seturile de date preluate și utilizate în cadrul acestui proiect de pe data.gov.ro sunt:

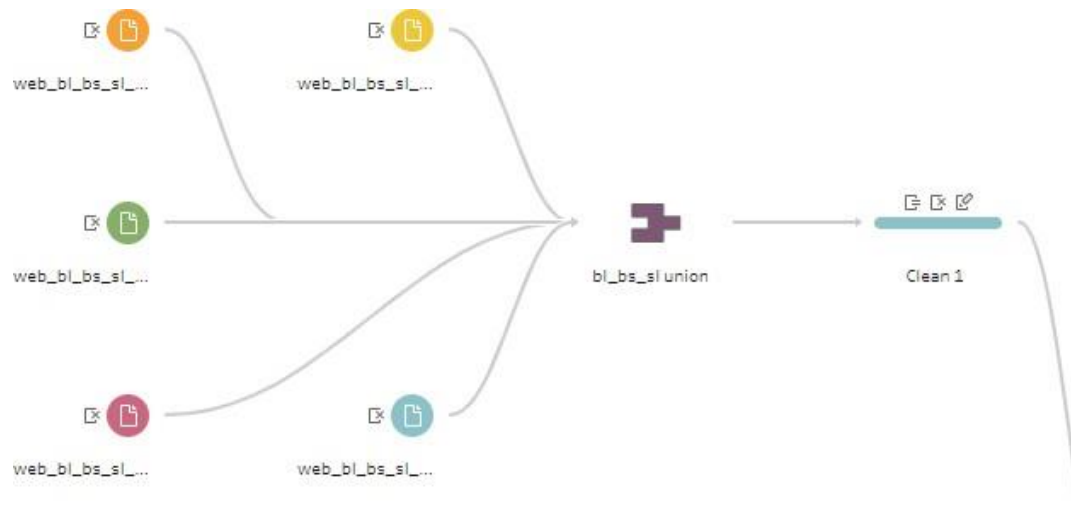
- *web\_uu\_an2018* format TXT și CSV
- *web\_uu\_an2019* format TXT și CSV
- *web\_uu\_an2020* format TXT și CSV
- *web\_uu\_an2021* format TXT și CSV
- *web\_uu\_an2022* format TXT și CSV
- *web\_ir\_an2018* format TXT și CSV
- *web\_ir\_an2019* format TXT și CSV
- *web\_ir\_an2020* format TXT și CSV
- *web\_ir\_an2021* format TXT și CSV
- *web\_ir\_an2022* format TXT și CSV
- *web\_bl\_bs\_sl\_an 2018* format TXT și CSV
- *web\_bl\_bs\_sl\_an 2019* format TXT și CSV
- *web\_bl\_bs\_sl\_an 2020* format TXT și CSV
- *web\_bl\_bs\_sl\_an 2021* format TXT și CSV
- *web\_bl\_bs\_sl\_an 2022* format TXT și CSV
- *3firme\_neradiate\_cu\_sediu\_2023-01-07* format CSV
- *1firme\_neradiate\_fara\_sediu\_2023-01-07* format CSV

În vederea optimizării procesului de curățare am efectuat artificii de a înlocui numele coloanelor în documentele TXT cu versiuni prescurtate ale acestora, forma fiind o înșiruire pe baza de indice pornind de la i1 și continuând până la i17. Denumirea completă a coloanelor a rămas în formatul original în documentele CSV. Figura1 de mai jos este o reprezentare vizuală a acestei convenții:

Creante;i4	
Casa si conturi la banci;i5	
CHELTUIELI IN AVANS;i6	
DATORII;i7	
VENITURI IN AVANS;i8	
PROVIZIOANE;i9	
CAPITALURI - TOTAL din care;;i10	
Capital subscris varsat;i11	
Patrimoniul regiei;i12	
Cifra de afaceri neta;i13	
VENITURI TOTALE;i14	
CHELTUIELI TOTALE;i15	
Profitul brut;i16	
Pierdere bruta;i17	

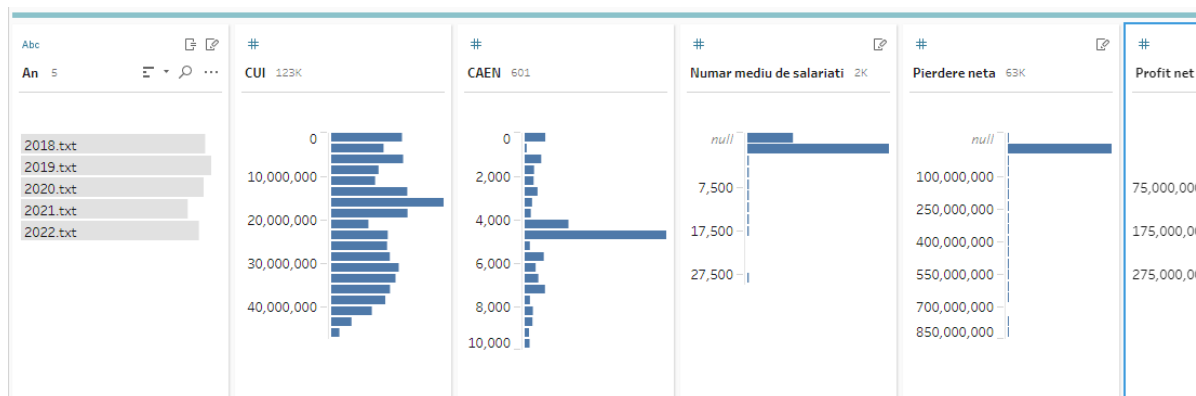
Figură 1

În urma analizării celor 5 ani preluați pentru perioada acestui studiu am constatat că numele coloanelor își păstrează aceeași formă în cadrul fișierelor de același tip; fapt pentru care la pasul următor vom realiza o uniune, continuat imediat de o etapă de curățare după cum se poate observa în Figura2 anexată mai jos. Rezultatul pasului de curățare s-a materializat prin modificarea numelui coloanelor ce prezentau obiect de interes cu denumirile preluate din documentul CSV.



Figură 2

Figura 3 prezintă vizual finalitatea pașilor amintiți anterior. Operațiunile efectuate în cadrul acestor procese au vizat fișierele web\_bl\_bs\_sl. În continuare am repetat același procedeu pentru documentele web\_uu și web\_ir.



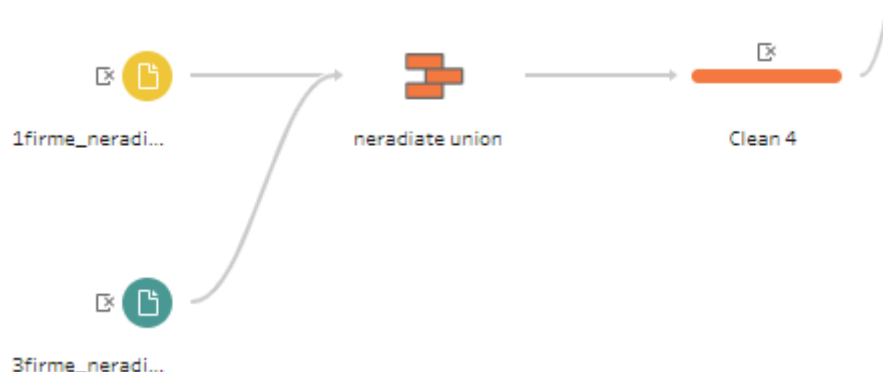
Figură 3

În etapa următoare am concatenat cele trei uniuni anterioare și am realizat un pas de curățare pentru a elimina coloanele irelevante pentru continuarea studiului nostru. Figura 4 expune o vizualizare a datelor pentru union-ul final al situațiilor financiare.

Abc	#	#	#	#	#	#	#
An	CUI	CAEN	Datorii	Cifra de afaceri	Venituri totale		CI
2018	15,469,928	111	0	0	0		n
2018	492,108	111	0	0	0		n
2018	9,719,248	111	0	0	0		n
2018	1,942,570	111	6,081	0	0		n
2018	1,308,875	111	25,203	0	0		n
2018	27,993,615	111	0	0	0		n
2018	12,683,624	111	0	0	0		n
2018	28,175,967	111	0	0	0		n
2018	2,124,714	111	0	0	0		n
2018	18,508,550	111	121	0	0		n
2018	7,965,475	111	0	0	0		n
2019	21,000,112	111	0	0	0		n

Figură 4

Ca urmare a existenței unui număr mare de înregistrări al firmelor studiate pentru perioada 2018/2022 am realizat procesul de excludere al celor inactive. Pentru a obține date relevante pentru studiul nostru vom selecta firmele ce se încadrează la nivel activ, din lista firmelor neradiate cu sau fără sediu, preluate de pe site-ul data.gov.ro. În urma descărcării datelor am realizat un pas de uniune a documentelor, iar cu ajutorul unei etape de curățare am păstrat doar datele importante analizei noastre.



Figură 5

Următorul pas a fost crearea unui Join între rezultatele fișierelor situațiilor financiare și cel al fișierelor cu informații despre neradierea firmelor pentru a rămâne doar cu cele active. Acest lucru l-am realizat pe baza valorii CUI-ului.

Applied Join Clauses				Clean 5		Clean 4	
Clean 5		=	Clean 4	CUI		CUI	
Join Type: Inner							
Click the graphic to change the join type.							
Clean 5							
Summary of Join Results							
Click the bar segments to view the included and excluded values.							
Mismatched values							
	Included		Excluded				
Clean 5	525,686		1,637,001				
Clean 4	393,270		0				
Join Result	525,686						

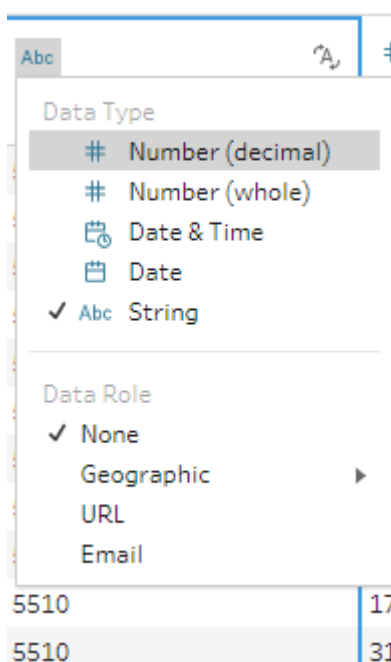
  

Clean 5	Clean 4
19	2,133
27	9,654
35	21,038
51	22,122
86	24,670
94	43,109
108	54,523
116	55,553
124	118,155
132	128,450
159	128,604
167	128,710
213	128,728
230	129,499
256	129,502
280	129,510

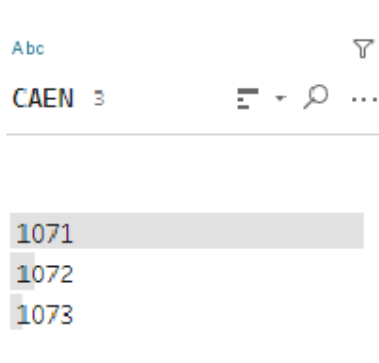
Figură 6

Pentru a nu obține coduri CAEN cu virgulă am efectuat conversia formatului coloanei asociate din Number(Whole) în String.

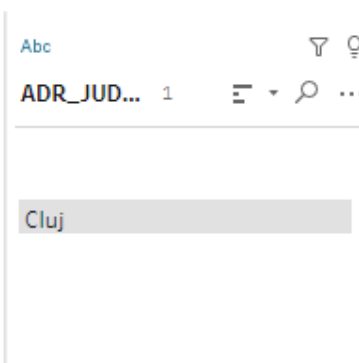
În cele din urmă am selectat doar codurile CAEN corespunzătoare fabricării produselor de brutărie și a produselor făinoase. Acestea au fost selectate utilizând Filter > Selected Values > Selecția propriu-zisă a codurilor dorite > Done. În plus, tot în cadrul acestui pas am selectat și valoarea adresei de județ ca fiind Cluj și am exclus datele nule din coloana ce conține informații despre numărul mediu de angajați.



Figură 7

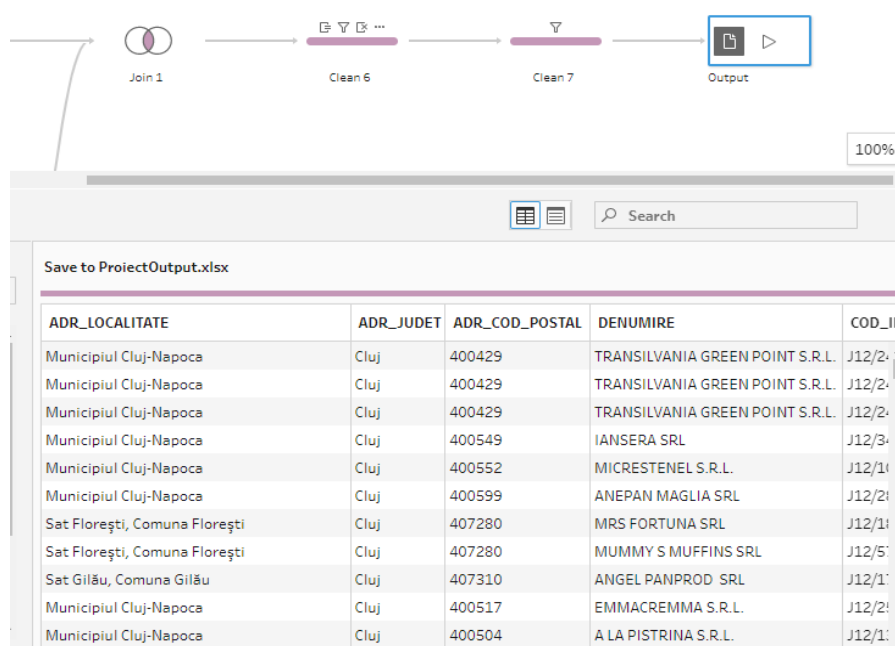


Figură 8



Figură 9

Produsul final se poate viziona creând un output, acesta ne va afișa datele obținute în urma tuturor proceselor de editare și sortare:



Figură 10



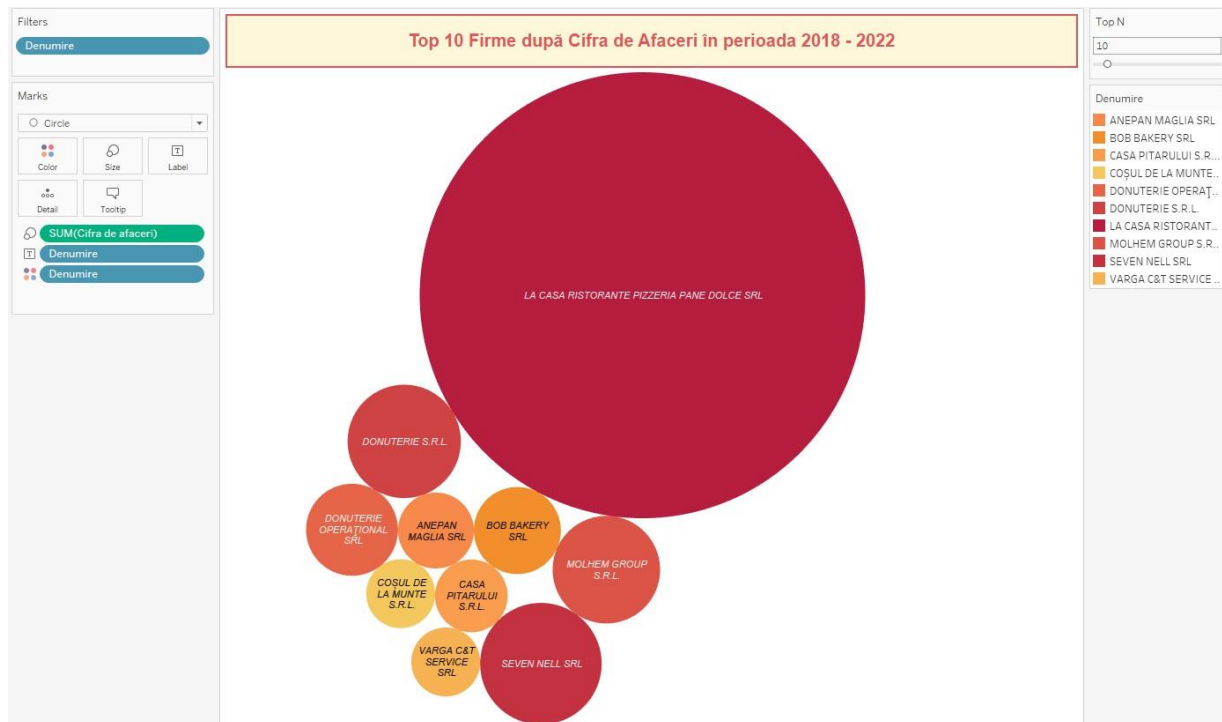
# Vizualizarea și Analiza Datelor

## Topul firmelor

În diagrama din Figura 11 se realizează reprezentarea vizuală a clasamentului firmelor cu cele mai mari cifre de afaceri înregistrate în perioada 2018/2022. Am aplicat Filtrarea după denumire și am setat următoarele note:

- Dimensiunea a fost setată în funcție de SUM(Cifra de afaceri)
- Eticheta a fost setată după denumire
- Culoarea a fost aleasă în funcție de Denumire

În partea dreaptă a diagramei este regăsită legenda alături de numărul de firme pe care dorim să îl regăsim în topul nostru.



Figură 11

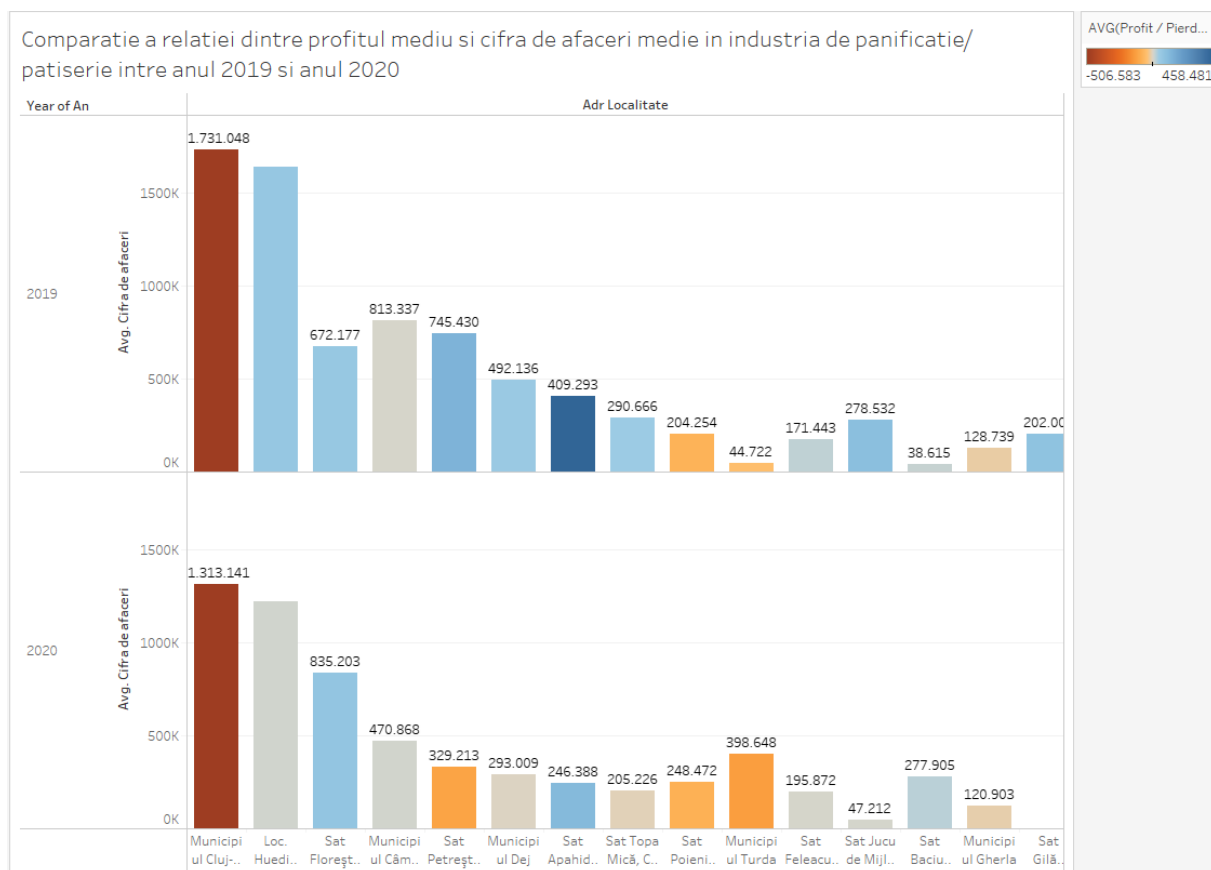
## Comparații

Este important ca în continuare să analizăm și să conturăm o comparație a relației dintre profitul mediu și cifra de afaceri medie pe piața industriei de produse de panificație și patiserie. Acest studiu ne va ajuta să vizualizăm situația pieței pentru perioada 2019/2020. În continuare am setat ca și filtru anul, iar ca note avem:

- Culoarea a fost setată în funcție de AVG(Profit / Pierdere)
- Eticheta a fost setată în baza AVG(Cifra de afaceri)

Ca și coloană avem Localitatea, iar pentru linii am setat anul și cifra de afaceri.

În partea dreaptă a graficului avem legenda distribuției culorilor în funcție de profit/pierderi.

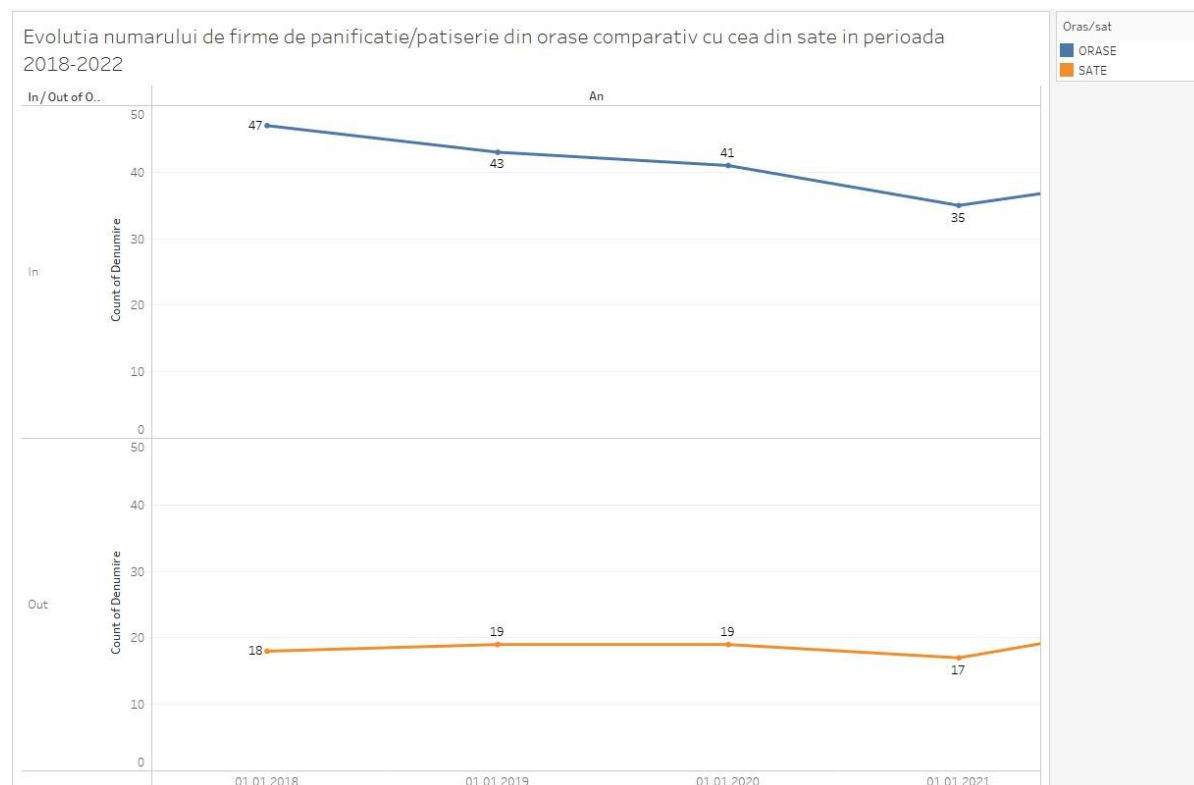


Figură 12

## Evoluția numărului de firme în funcție de mediu

În diagrama din Figura 13 am surprins evoluția numărului de firme cu specific de panificație/patiserie atât din mediul urban, cât și cel rural. Prin acest studiu putem crea o paralelă vizuală în expansiunea industriei făinoase, perioada de analiză fiind 2018/2022.

Graficul superior încadrează numărul de firme existente la oraș pentru fiecare an prezentat, în subsolul său fiind repartizat un grafic similar pentru situația de la sate. Ca și sistem de referință a fost preluată data de 1 Ianuarie pentru fiecare an.



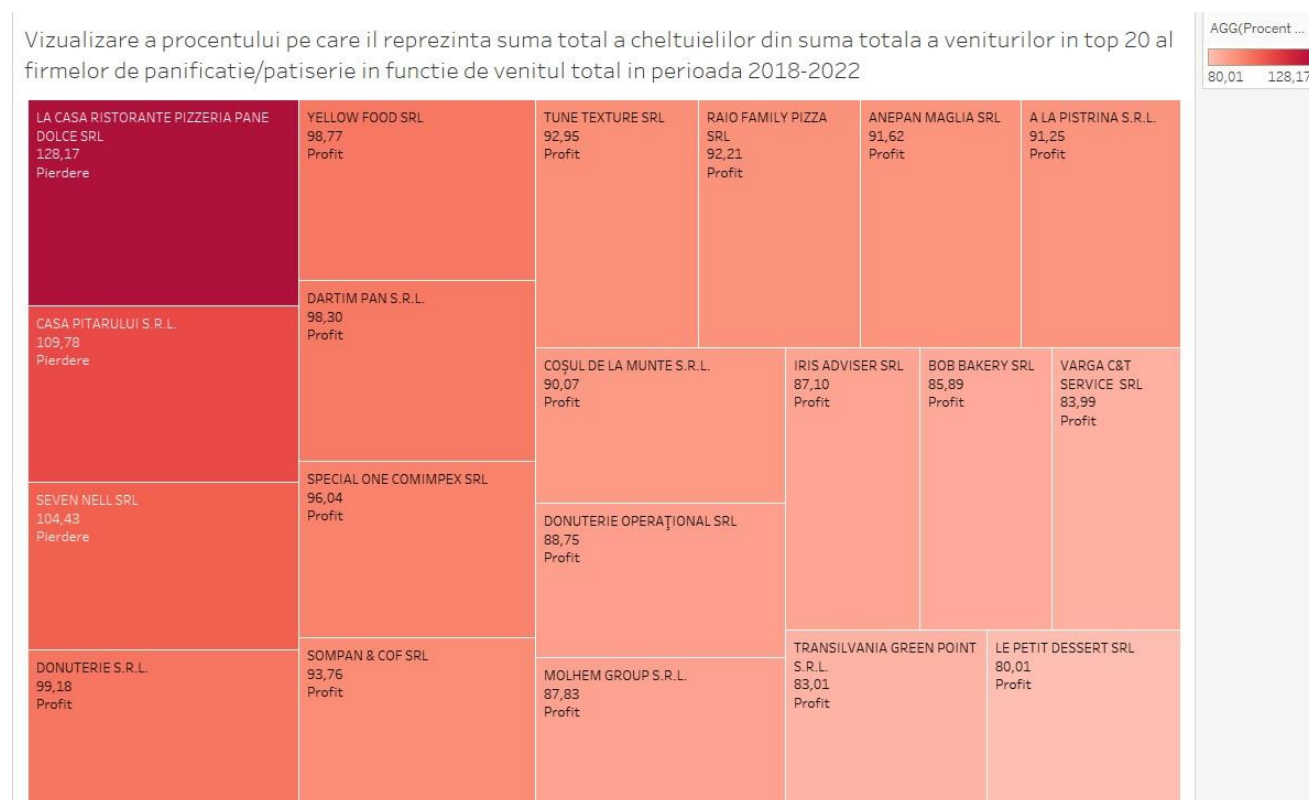
Figură 13

## Procent Cheltuieli din Venituri

În analiza de business am considerat relevant să vizualizăm procentele reprezentate de suma totală a cheltuielilor din suma totală a veniturilor, de data aceasta luând ca referință top 20 de firme pentru specificul ales. În vederea acestui studiu am luat ca și referință perioada 2018/2022 pentru a putea identifica și evidenția în mod eficient firmele care se situează pe profit, respectiv pierdere.

Pentru vizualizare avem legenda de AVG(Procent cheltuieli din venituri), am setat filtru în funcție de denumire și am aplicat următoarele note:

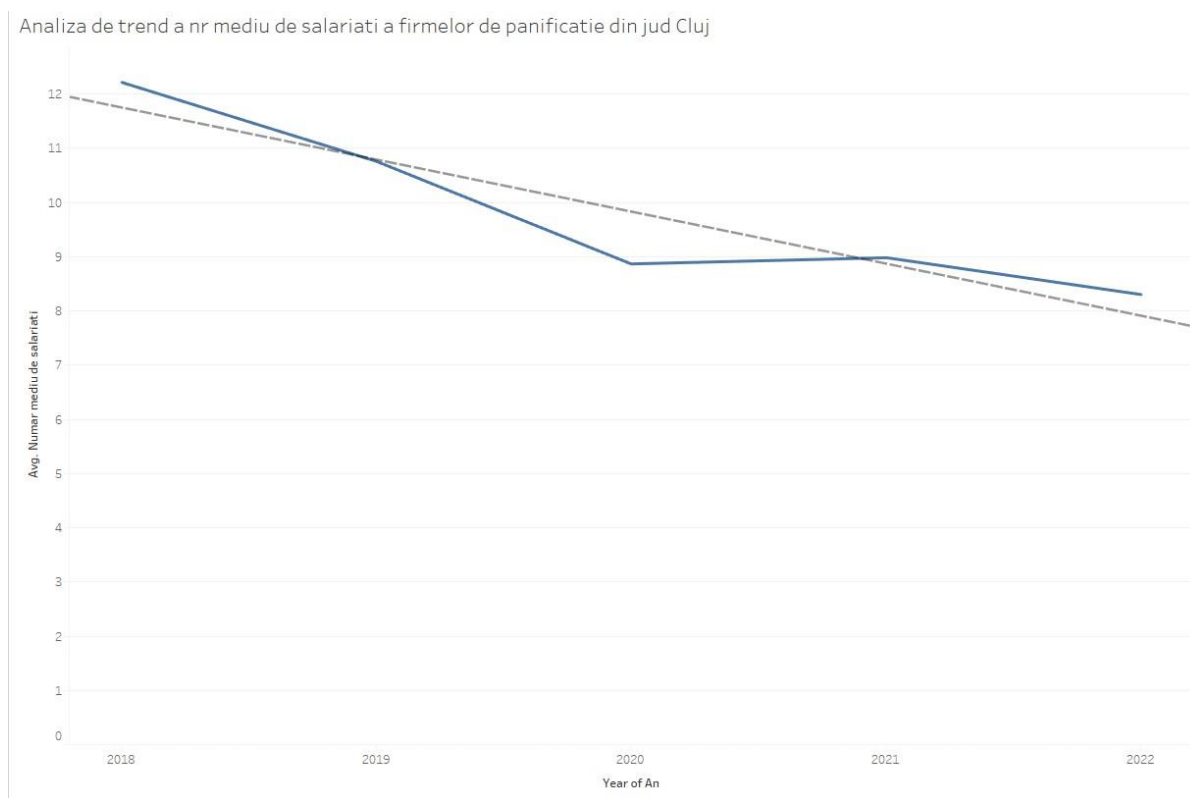
- Culoarea a fost repartizată în funcție de AVG(Procent cheltuieli din venituri)
- Dimensiunea a fost setată după considerentul AVG(Procent cheltuieli din venituri)
- Eticheta 1 a fost alocată pentru Denumire
- Eticheta 2 a fost alocată în funcție de AVG(Procent cheltuieli din venituri)
- Eticheta 3 a fost setată în funcție de AVG(Profit/Pierdere în funcție de procent cheltuieli din venituri)



Figură 14

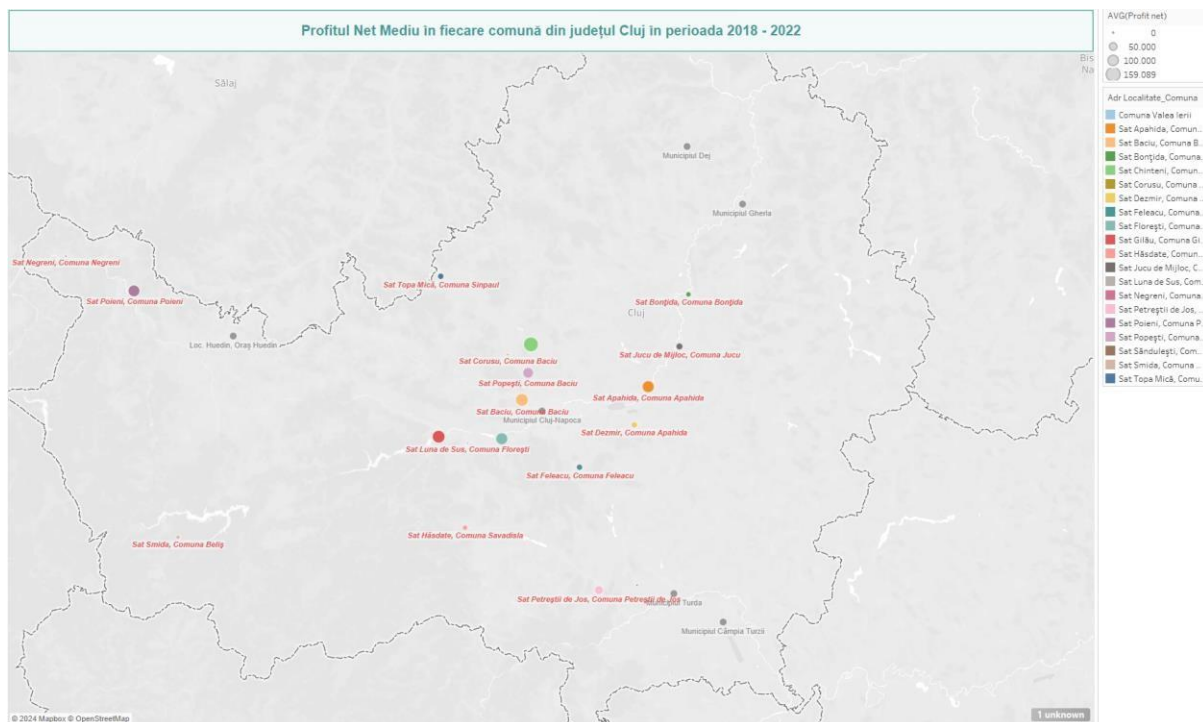
## Analiza de trend pentru numărul de salariați

Diagrama din Figura15 surprinde numărul mediu de salariați al firmelor de panificație din județul Cluj, pe o perioadă de 5 ani. Putem observa că graficul îmbrățișează un trend descendent în baza acestei vizualizări. Ca și coloană am selectat anii, iar ca și linie avem AVG(Număr mediu de salariați).



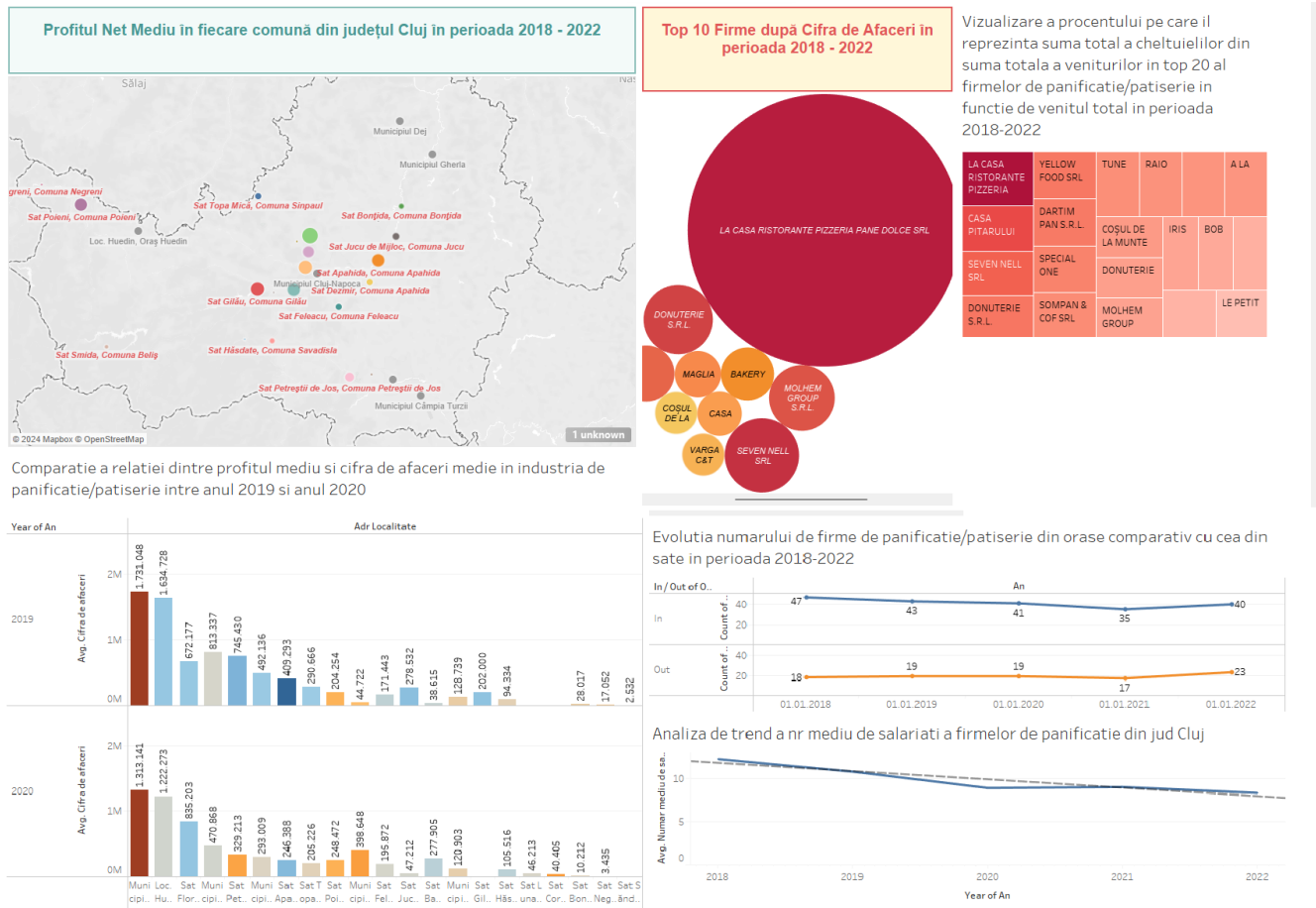
Figură 15

# Harta



Figură 16

Dashboard



## Story

Ca și încheiere am ales să surprindem cele mai importante caracteristici identificate în acest proiect printr-un panel de tip Story. Acesta surprinde 4 elemente cheie în analiza noastră și facilitează accentuarea trăsăturilor vitale analizei.

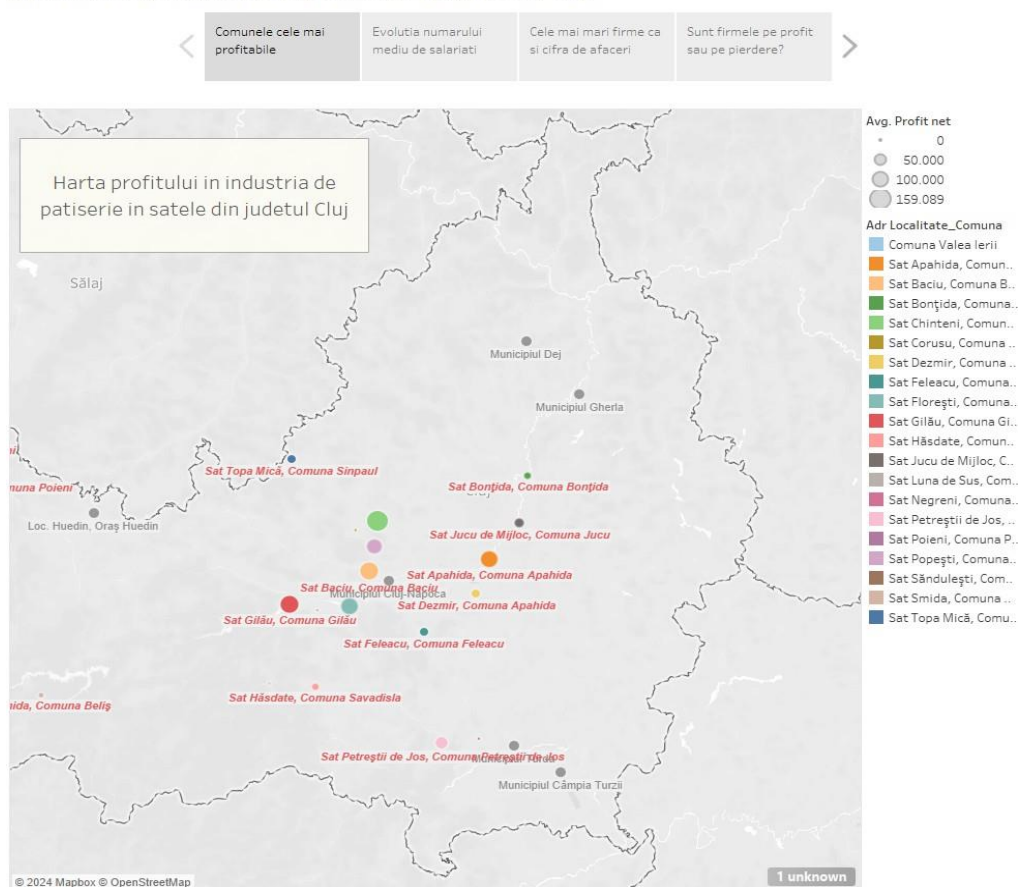
Prima parte surprinde Harta profitului în industria de patiserie în satele din județul Cluj, dimensiunea pe hartă fiind stabilită în baza profitului net, iar identificatorul de culoare a fost creat în baza localității. Harta surprinde totodată și legenda pentru fiecare câmp încadrat.

Elementul următor în poveste surprinde trendul descendent în ceea ce privește numărul mediu de angajați într-un magazin cu specific de panificație, de-a lungul a 5 ani.

Pentru o analiză de business este esențial să putem identifica liderul de piață stabilit în funcție cifra de afaceri, și dimensiunea acestuia în comparație cu brand-urile concurente. Putem observa astfel, că La Casa Ristorante Pizzeria Pane Dolce SRL se clasează în topul clasamentului.

Ca și etapă finală vom vizualiza situația profitului și a pierderilor, repartizate în funcție de județ pentru industria de produse de panificație și patiserie. Surprindem astfel prin acest grafic o privire de ansamblu asupra evoluției pieței de la un an la altul, în situația de față raportându-ne la perioada 2019/2020.

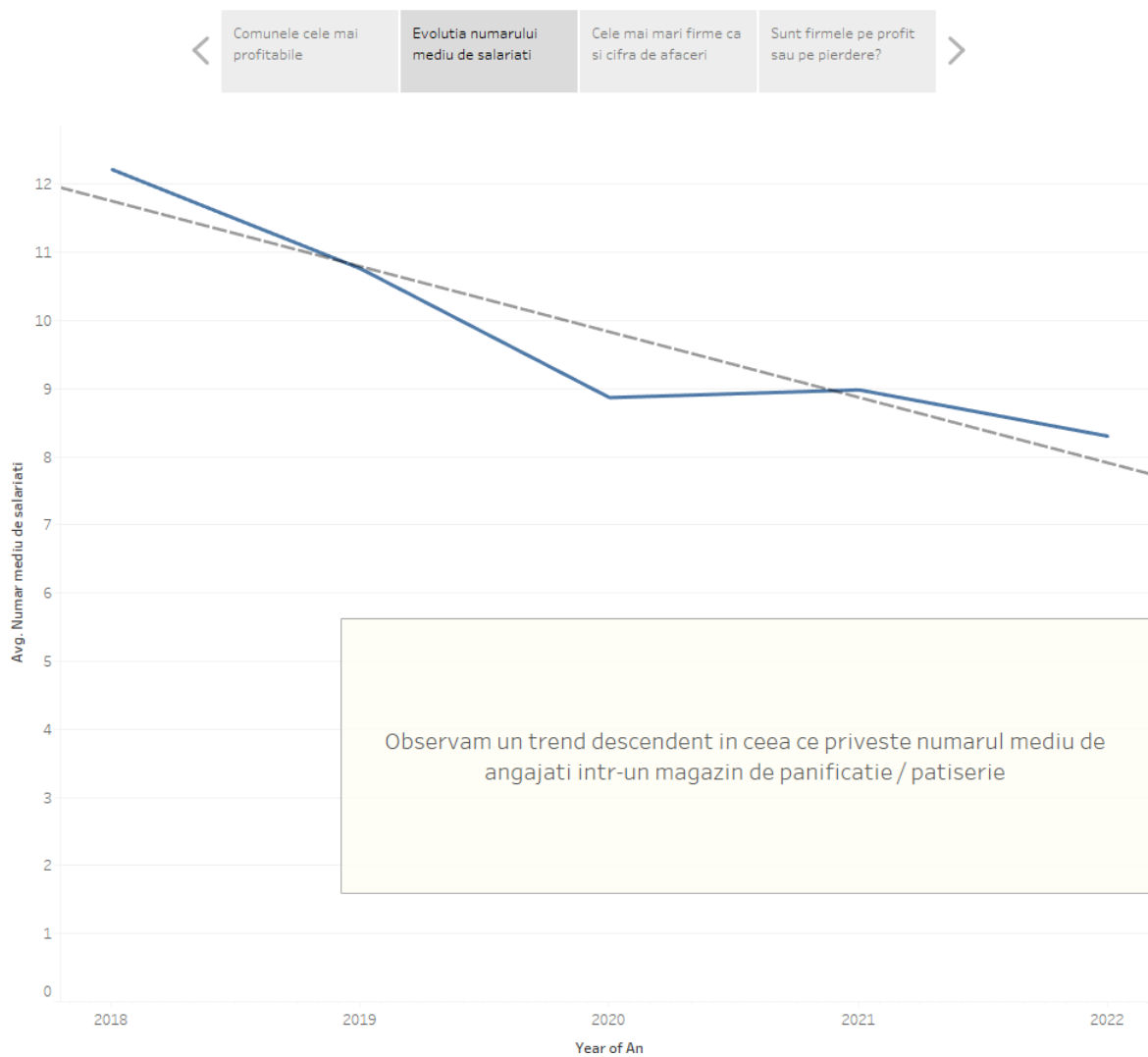
### Industria de panificatie si patiserie din judetul Cluj



Figură 18

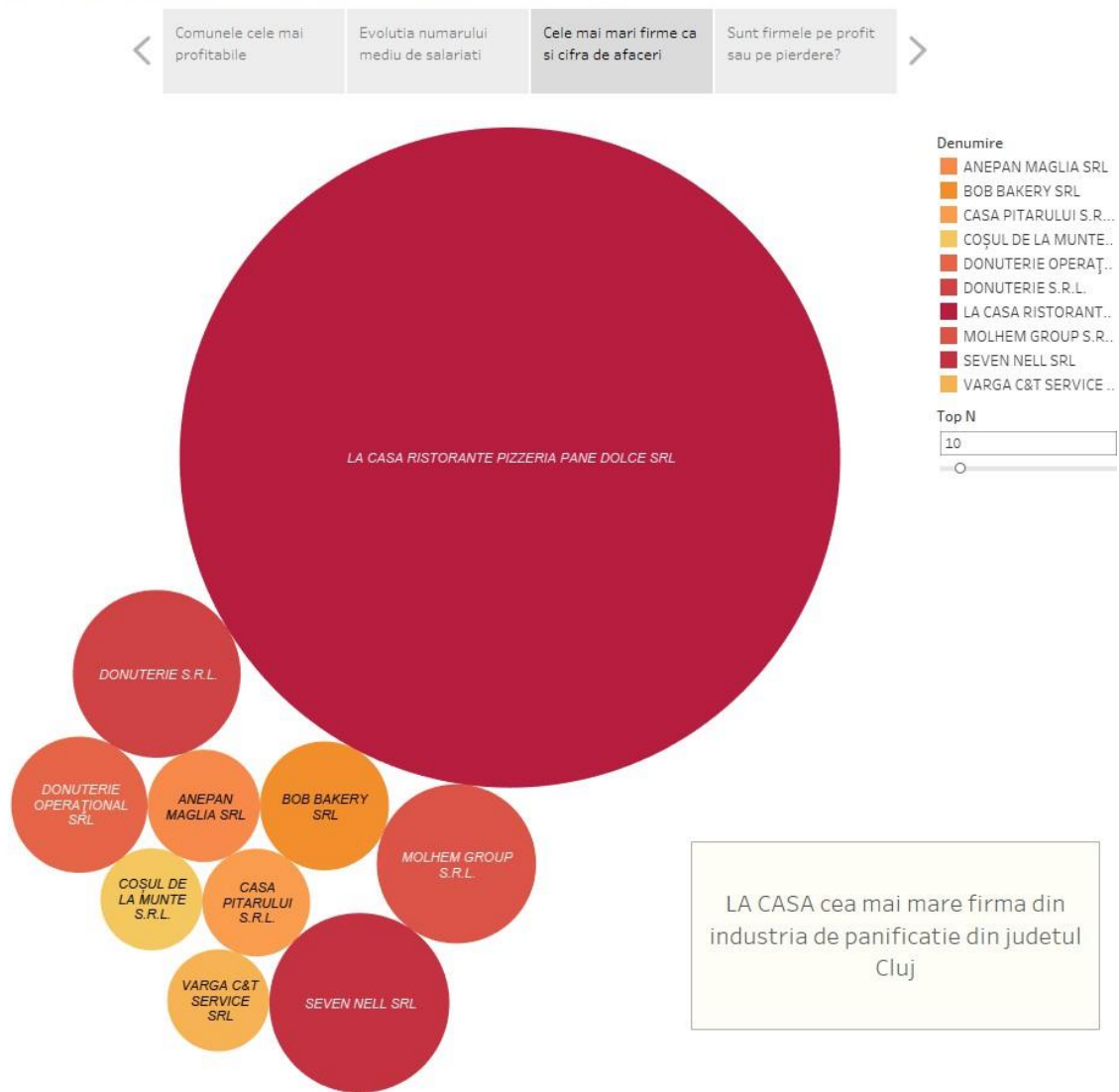


## Industria de panificatie si patiserie din judetul Cluj



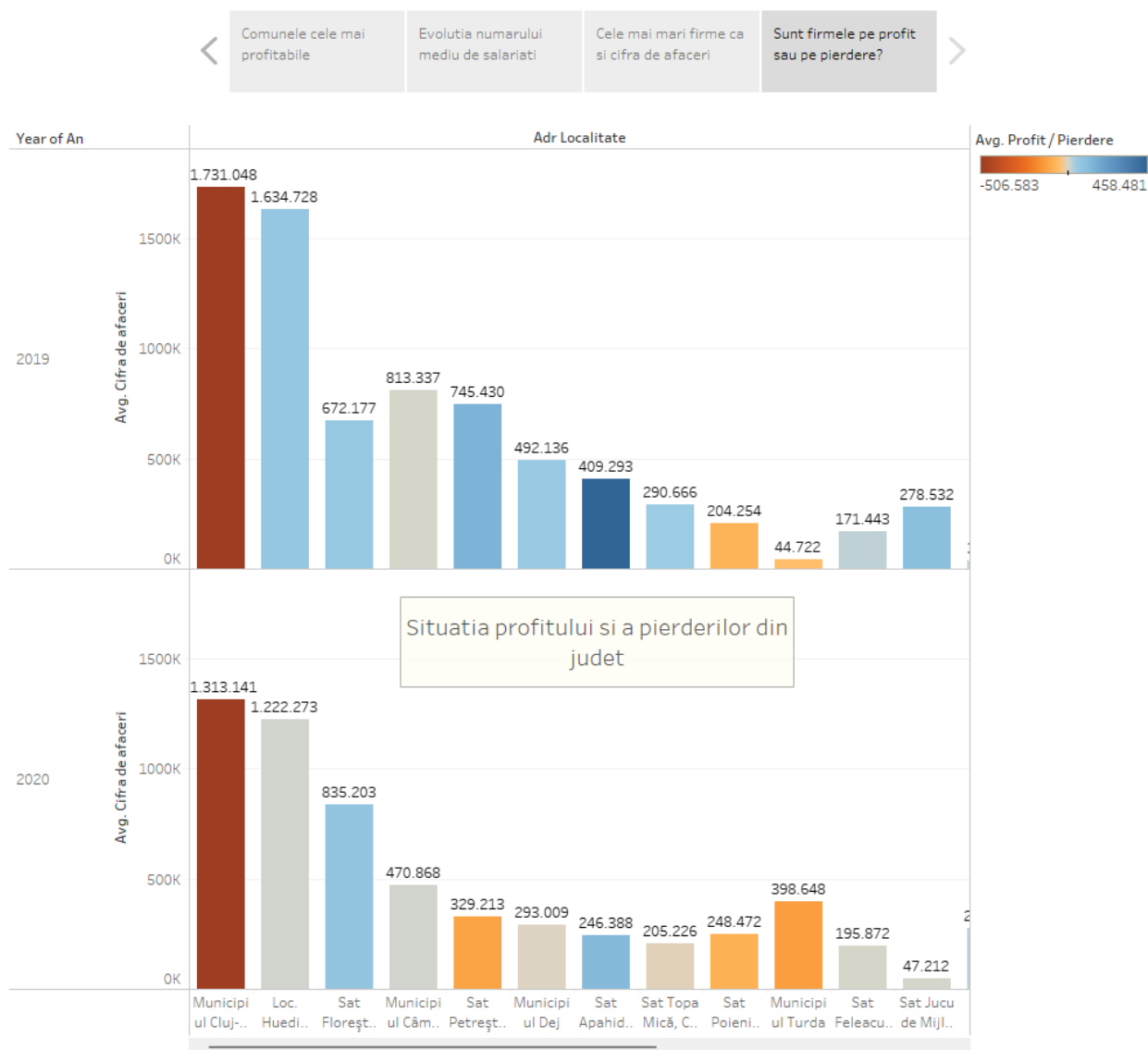
Figură 19

Industria de panificatie si patiserie din judetul Cluj



Figură 20

## Industria de panificatie si patiserie din judetul Cluj



Figură 21