

Proiect Business Intelligence

Acest proiect a fost realizat de:
Petoaca-Nelega Vlad-Natalian
Stavnic Adrian
Cordovan Raul
Tonko Zsanett-Julia-Vivien

1.Descriere problema de Business pentru care s-a realizat proiectul

In cadrul proiectului am analizat situatiile financiare a firmelor ce au ca obiect de activitate turismul si serviciile de cazare pe raza judetului Cluj. Pentru a realiza acest lucru am folosit urmatoarele coduri CAEN:

Hoteluri 55

5510 - Hoteluri si alte facilități de cazare similare

5520 - Facilități de cazare pentru vacante si perioade de scurta durata

5530 - Parcuri pentru rulote, campinguri si tabere

5590 - Alte servicii de cazare

Factorii analizati pentru fiecare firma sunt urimatorii:

- Cifra de afaceri neta
- Veniturile totale
- Cheltuielile totale
- Profitul net
- Pierdere neta
- Pierdere bruta
- Profitul brut
- Numar mediu de angajati

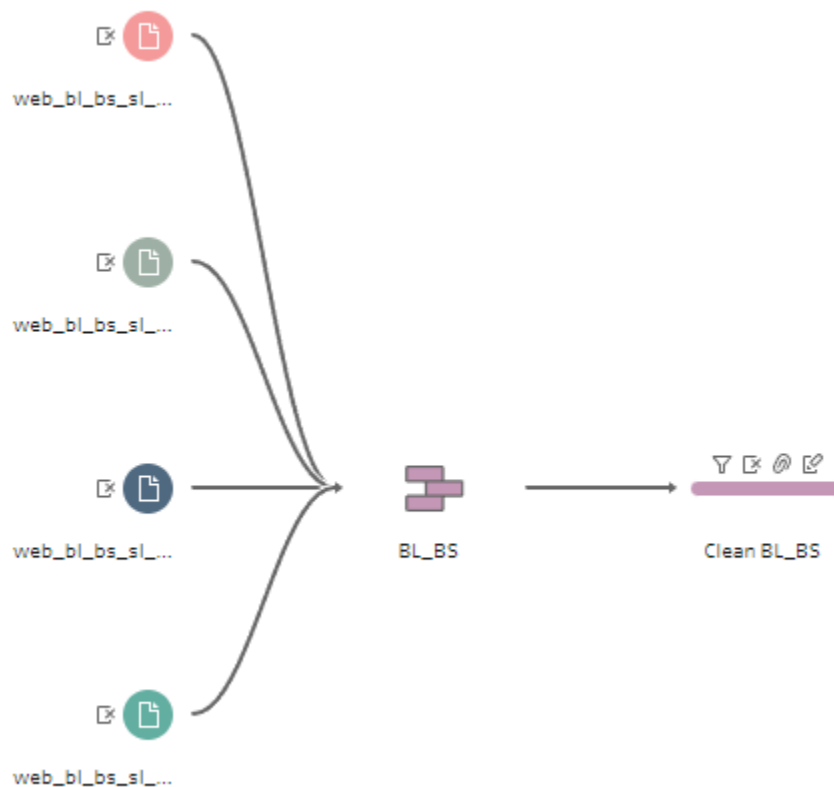
Acesti factori ne-au permis sa evaluam in detaliu performantele financiare ale firmelor din acest sector si sa obtinem o imagine clara a dinamicii economice din industria turismului si a serviciilor de cazare din judetul Cluj.

2. Pregatirea datelor

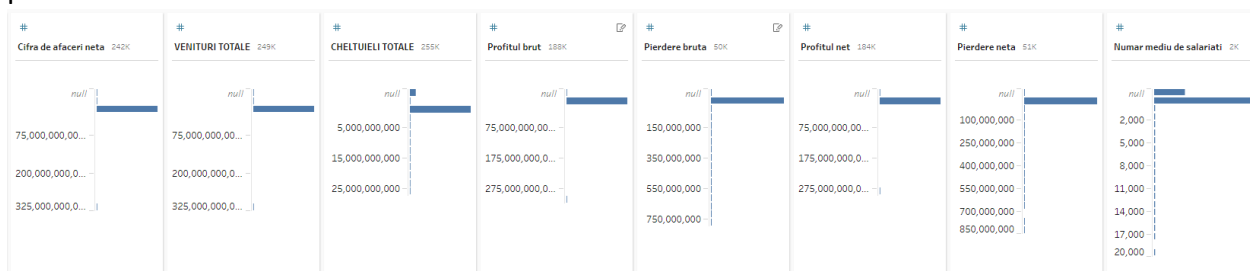
Am preluat datele de pe data.gov.ro de la sectiunea Ministerul Finantelor Publice ce contine Situatii financiare 2022, situatii financiare 2021 de unde am preluat datele pentru 2021, 2020 si 2019, analiza fiind facuta pe cei 4 ani consecutivi.

Am descarcat fisierele cu numele similar WEB_IR_AN si anul respectiv, WEB_BL_BS_SL_AN si anul respectiv si WEB_UU_AN si anul respectiv.

Am sectionat fisierele in 3 grupuri de tip Union de cate 4 fisiere. Am procedat la fel pentru toate fisierele, apoi de la fiecare Union a fost creat un Clean.



In acest clean am modificat numele sectiunilor ce in capul fisierele excel descarcat de pe data.gov, cu rolul de a fi sablon cu numele sectiunilor marcate pentru fisierele cu date de tip .txt. Am avut nevoie de sectiunile de la i13 pana la i20, redenumite ca si Cifra de afaceri neta; i13, VENITURI TOTALE; i14, CHELTUIELI TOTALE; i15, Profitul brut; i16, Pierdere bruta; i17, Profitul net; i18, Pierdere neta; i19, Numar mediu de salariati; i20, apoi CUI si CAEN care erau prestabilite.



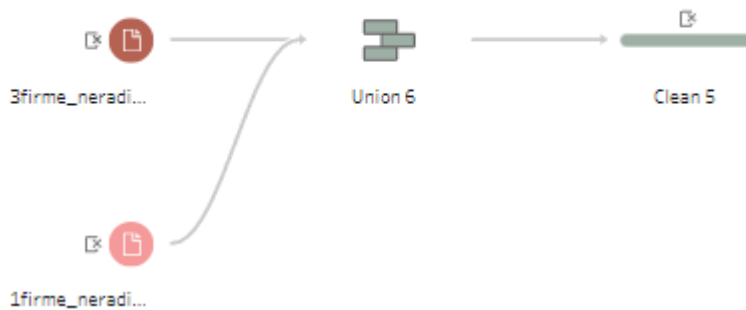
Am modificat numele cadranului cu numele fisierele in "An" si am modificat numele fisierele WEB sa fie denumite doar cu anul lor si formatul .txt.

Am unit toate cele 3 Clean-uri intr-un Union care mergea mai apoi la un Clean.

In Clean am adaugat cate un filtru pentru fiecare sectiune pentru a scoate valorile null.

Tot de pe site-ul Data.gov.ro am descarcat din pagina “Firme înregistrate la Registrul Comerțului până la data de 07 ianuarie 2023” documentele de format .csv ce detineau detalii despre firmele neradiate cu sediu si fara sediu, respectiv documentul 1 si 4.

Le-am introdus si pe aceste in proiectul nostru, apoi le-am conectat intr-un Union. De la Union a plecat un Clean din care am facut niste modificari. Am exclus cateva tabele.



Dupa Clean a fost legat un Join, ce conecta cleanul fisierelor interconectate de la inceput cu documentele de acum. Acest Join a fost urmat de un Clean. Deoarece am avut tabelul similar CUI la ambele grupuri a trebuit sa stergem unul din acestea.



Am schimbat CAEN SI CUI sa fie format String. Am dat split la Cod inmatriculare pentru a face rost de codul judetului care pentru orasul Cluj a fost J12, iar restul tabelelor au fost sterse.

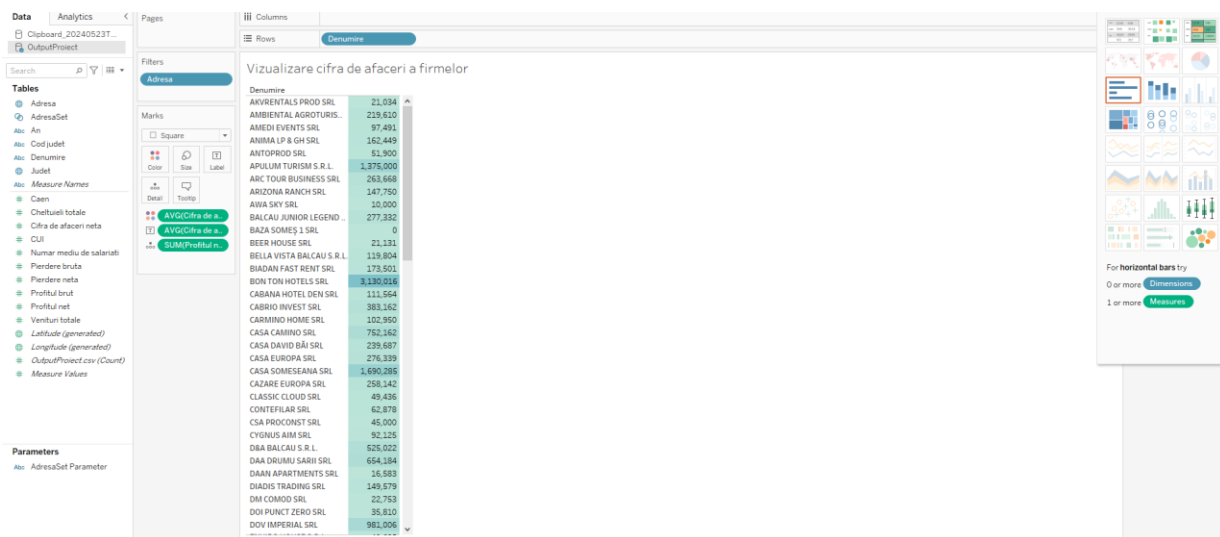
In clean 7 am schimbat numele cadranelor si am filtrat cadranul CAEN sa arate codurile 5510 - Hoteluri si alte facilități de cazare similare, 5520 - Facilități de cazare pentru vacante si perioade de scurta durata, 5530 - Parcuri pentru rulote, campinguri si tabere, 5590 - Alte servicii de cazare.

3. Data Visualization & Analysis

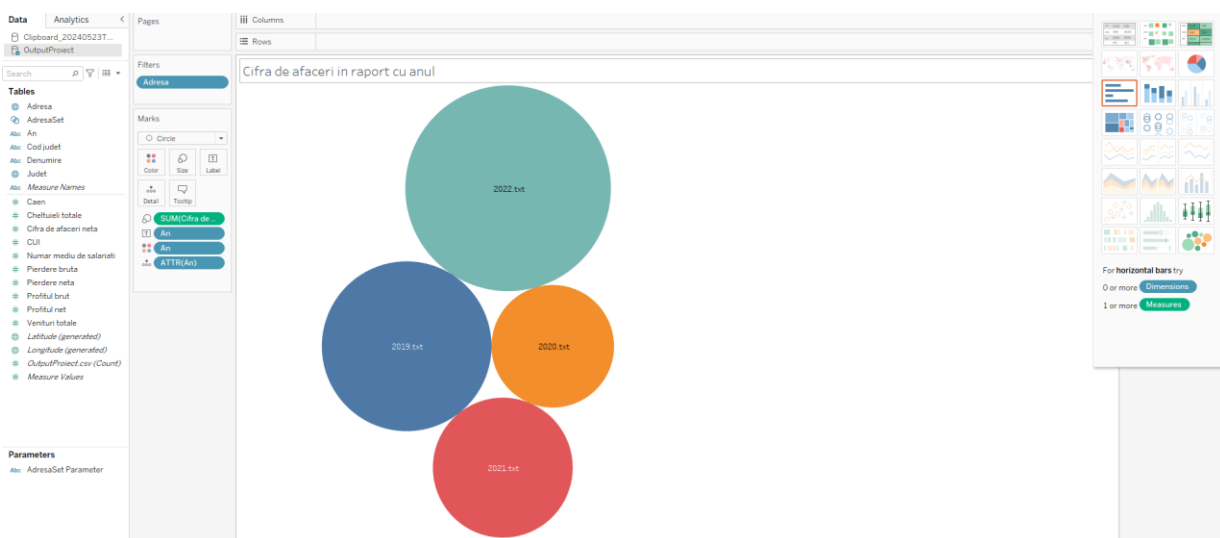
An	CUI	CAEN	Cifra de afaceri neta	Venituri totale	Cheltuieli totale	Profitul net	Pierdere neta	Pierdere bruta	Profitul brut	Numar mediu de salariati	Adresa	Judet	Denumire	Cod judet
2020.txt	33606284	5520	59,649	137,626	72,346	63,577	0	0	65,280	1	Sat Bonțida, Comuna Bonțida	Cluj	BONTIDA REST RESORT SRL	J12
2020.txt	34834272	5520	2,900	7,636	16,196	0	8,647	8,560	0	1	Sat Bucea, Comuna Negreni	Cluj	IRANAA ZEN SRL	J12
2020.txt	35018680	5520	5,714	11,940	140,640	0	128,757	128,700	0	1	Sat Rogojel, Comuna Săculeu	Cluj	POIANA VLĂDEȘI SRL	J12
2020.txt	36000265	5520	102,203	108,012	45,682	61,346	0	0	62,330	2	Municipiul Cluj-Napoca	Cluj	DIADIS TRADING SRL	J12
2020.txt	34926660	5520	133,138	145,438	167,942	0	23,764	22,504	0	2	Sat Săndulești, Comuna Săndulești	Cluj	HAPPY LOLA SRL	J12
2020.txt	35871600	5520	36,478	65,589	60,969	4,272	0	0	4,620	2	Sat Băbuțiu, Comuna Vultureni	Cluj	PETRINJELUL SRL	J12
2020.txt	35879248	5520	250,806	313,726	370,910	0	59,479	57,184	0	5	Sat Urca, Comuna Vișoara	Cluj	ANTOLECSIA HOUSE S.R.L.	J12
2020.txt	35346059	5590	59,083	59,083	57,170	546	0	0	1,913	1	Municipiul Cluj-Napoca	Cluj	PANNONIA GOLD SRL	J12
2020.txt	35844053	5590	72,289	79,322	64,965	13,961	0	0	14,357	1	Municipiul Cluj-Napoca	Cluj	GUIDER TECH SRL	J12
2020.txt	34388323	5590	1,286	1,286	3,877	0	2,591	2,591	0	1	Municipiul Cluj-Napoca	Cluj	DM COMOD SRL	J12
2020.txt	25822988	5590	380,806	381,010	374,451	3,131	0	0	6,559	1	Municipiul Dej	Cluj	ARC TOUR BUSINESS SRL	J12
2020.txt	35581763	5520	38,860	38,860	26,726	11,214	0	0	12,134	1	Sat Luna de Sus, Comuna Florești	Cluj	BALINTGAZDA SRL	J12

In Tableau Desktop am importat fisierul csv exportat din proiectul anterior.

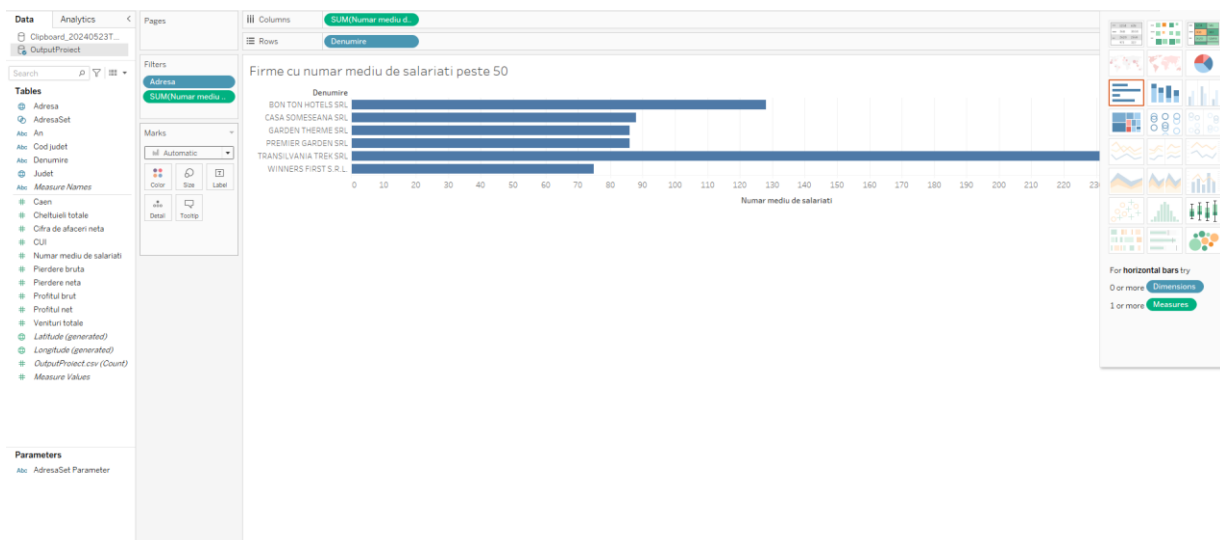
Prima pagina reprezinta vizualizarea cifrei de afaceri a firmelor:



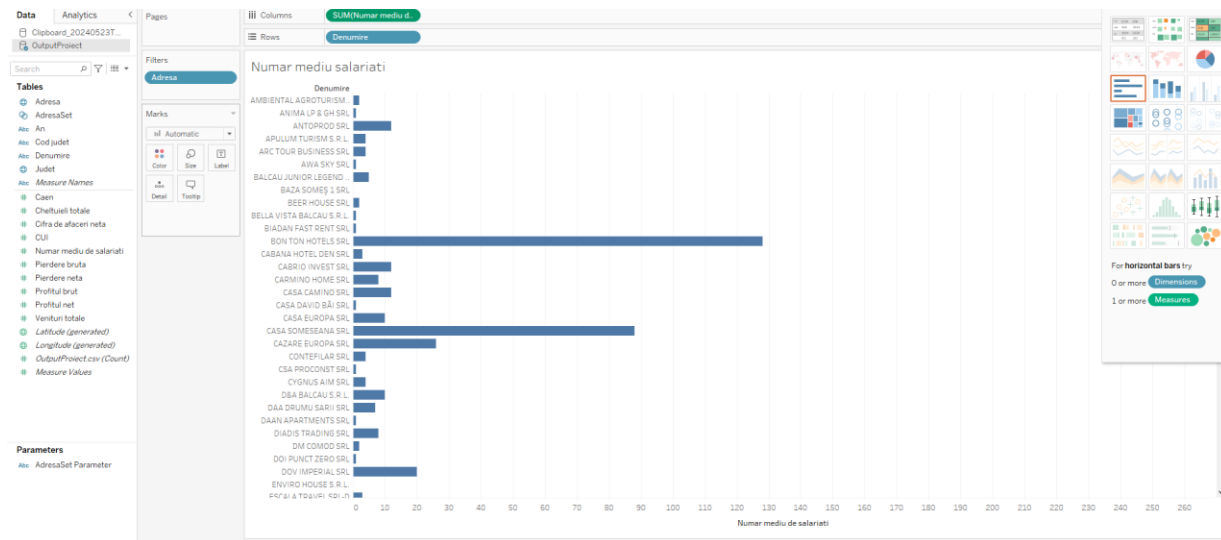
Aceasta pagina descrie cifra de afaceri in raportul cu anul specific:



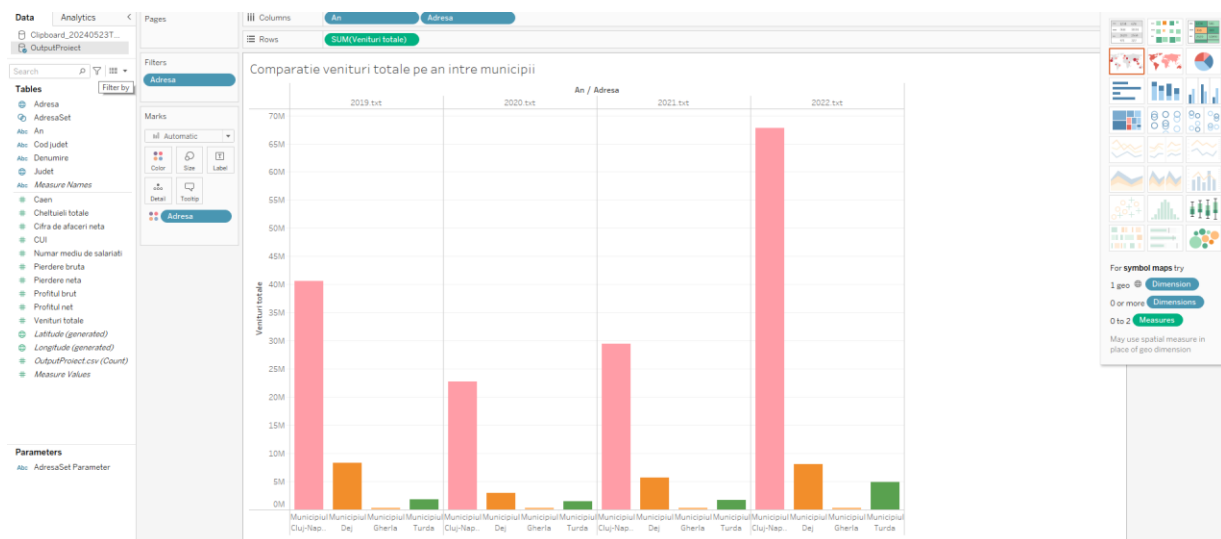
Pagina aceasta exemplifica grafic firmele cu numar mediu de salarizati peste 50:



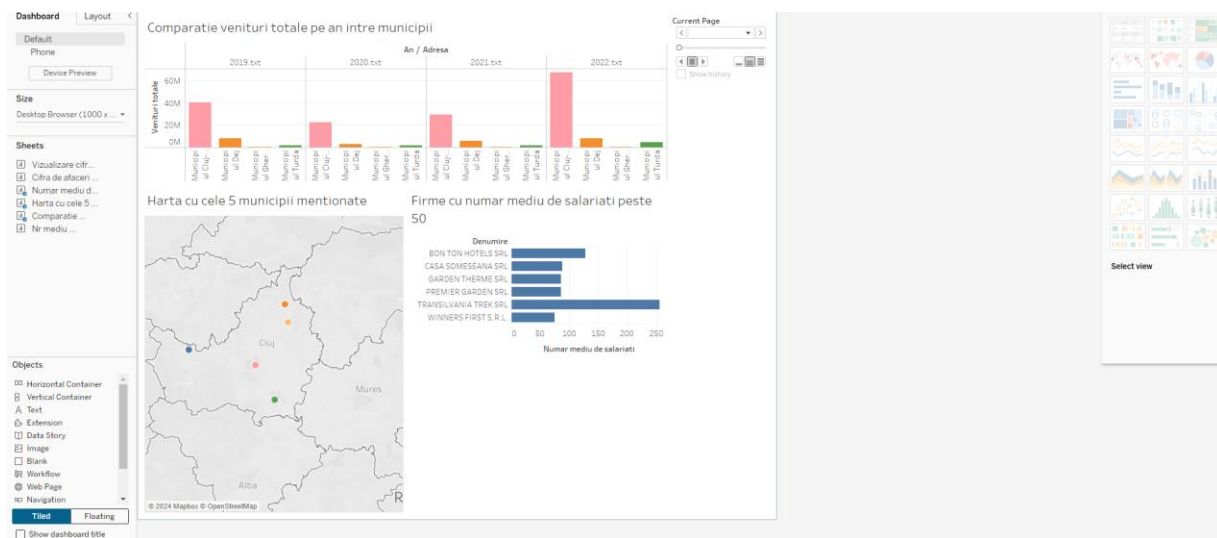
Acest grafic arata numarul mediu de salarizati:



Acest tabel descrie Comparatia veniturilor totale pe an intre municipii



Aceasta contine 3 din cele mai importante tabele si facut in format Dashboard:



Story-ul l-am facut pe baza tabelului cu salariatii medii cu valoarea peste 50.

