Analiza domeniilor de construcții și

transporturi de mărfuri pe

sectorul ALBA

1. **DESCRIERE PROBLEMĂ DE BUSINESS**

Piaţa construcţiilor, reprezentată de domeniul de activitate cu codul CAEN **4120**, este destul de echilibrată. Cele mai importante companii din industria de construcţii contractează proiecte mari de infrastructură, iar companiile medii şi mici care se ocupă de construcţii de ansambluri rezidenţiale, clădiri de birouri sau alte construcţii civile, însumează o valoare importantă din total.

De asemenea, o altă piață relevantă din acest punct de vedere este reprezentată de transporturile rutiere de mărfuri cu codul CAEN **4941** asociat. Se dorește o analiza acestor 2 sectoare în paralele pentru a reuși să conturăm o imagine de ansamblu asupra celui mai relevant domeniu în care o firmă aflată în stadiul deschiderii unei noi afaceri să poată vedea concret evoluția în timp a acestor domenii.

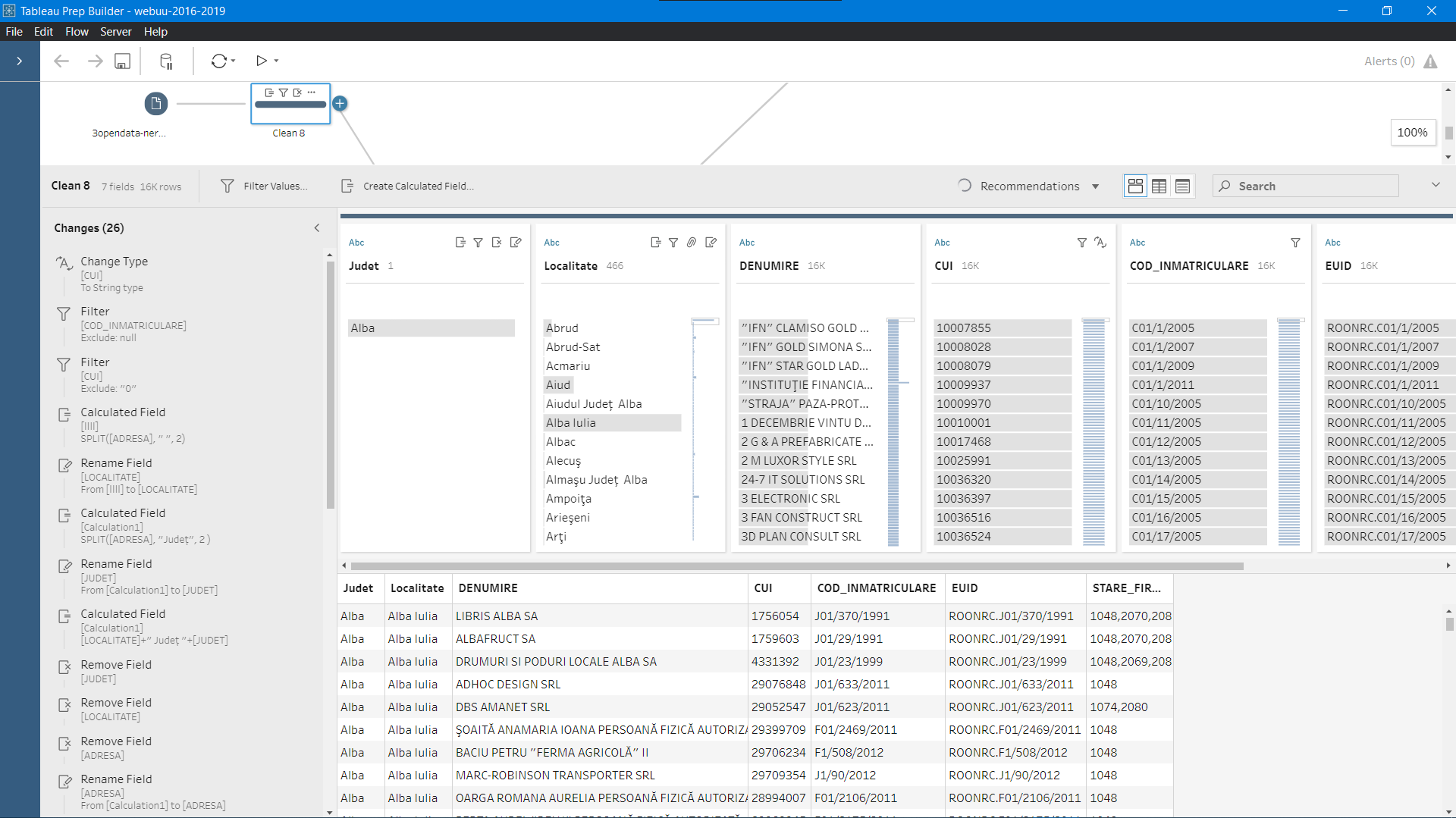
Ca să avem o imagine a structurii cred că este relevant să ştim că primele 10 companii din județul Alba care au o pondere mare în totalul cifrei de afaceri. Astfel, datele folosite în acest context au fost furnizate de pe site-ul: <https://data.gov.ro/dataset> pentru 4 ani consecutivi:

* Anul 2016 – webuuan2016.txt <https://data.gov.ro/dataset/situatii_financiare_2016/resource/7f631a71-8e41-4319-9a16-1129b02fb366>
* Anul 2017 - webuuan2017.txt <https://data.gov.ro/dataset/situatii_financiare_2017/resource/2c33ebbb-229c-427a-9752-9032a61e9613>
* Anul 2018 – webuuan2018.txt <https://data.gov.ro/dataset/situatii_financiare_2018/resource/86aa042a-8e42-4684-8d3f-971d849eda48>
* Anul 2019 – web\_uu\_an2019.txt <https://data.gov.ro/dataset/situatii_financiare_2019/resource/09a63c61-b5bb-448f-ae80-fb5aaf0a175a>

Datele din aceste fișiere au fost prelucrate în Tableau Prep Builder, curățate și unite cu datele respective firmelor de pe site-ul: <https://data.gov.ro/dataset?_organization_limit=0&organization=onrc>

1. **CURĂȚAREA DATELOR**

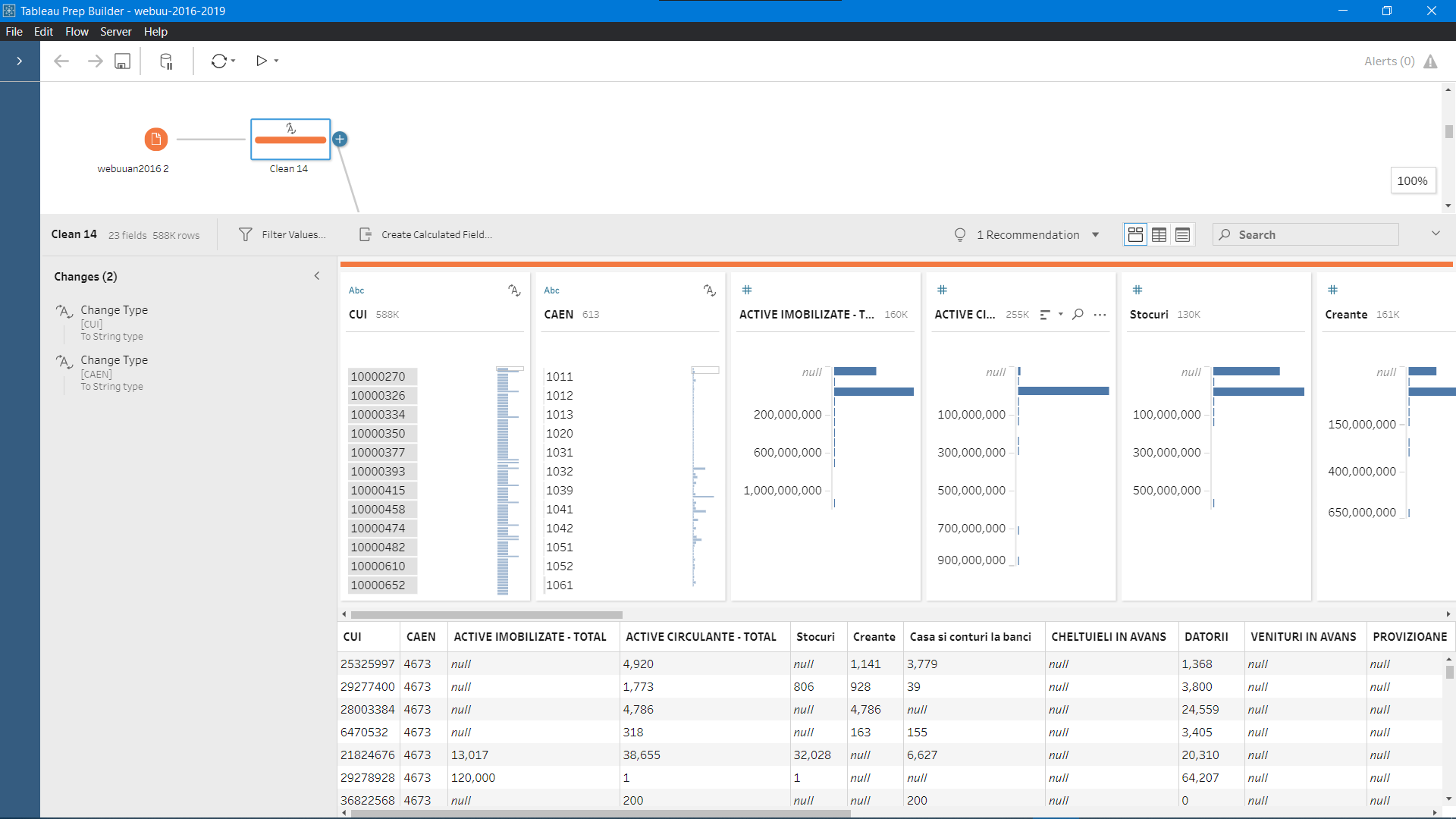
Pasul 1 a fost reprezentat de conectarea la setul de date descris anterior și curățarea datelor în Tableau Prep. Datele au necesitat modificări serioase în ceea ce ține de structurarea adresei deoarece de la an la an aceasta a fost înregistrată în moduri diferite, iar combinarea lor nu este posibilă decât dacă structura corespunde în fiecare fișier asupra căruia se aplică union datelor. Toate aceste modificări pot fi revăzute în cadrul figurii de mai jos.



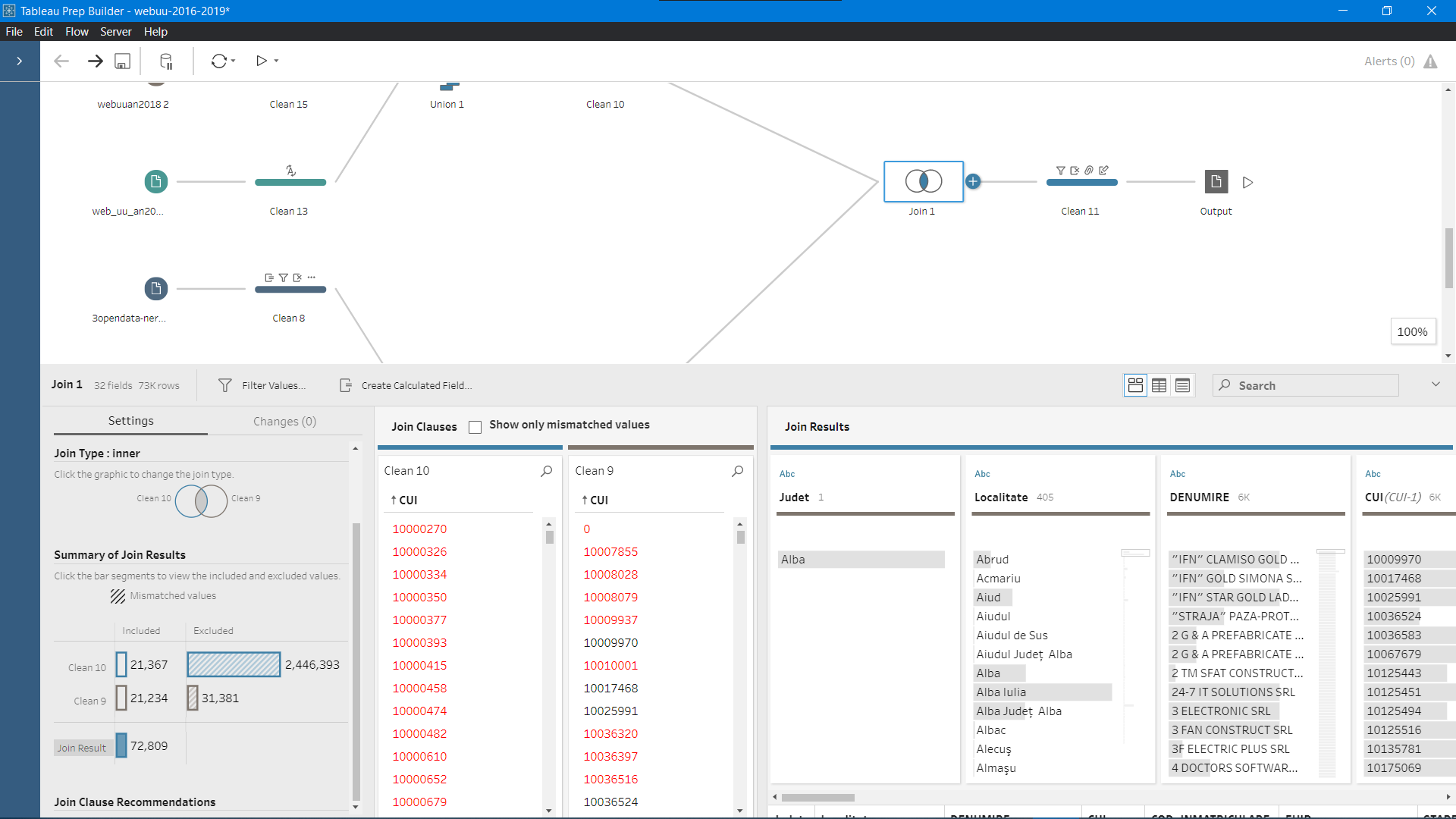
Datele au fost separate cu ajutorul formulei: ***SPLIT([ADRESA], " ", 2)*** iar mai apoi fiecare cîmp rezultat a fost supus eliminării spațiilor în plus, păstrând strict orașul și județul corespunzător. Excepție a făcut unul dintre fișiere a cărui structură era mult mai diferită și a necesitat și alte modificări.

În ceea ce privește înregistrările corespunzătoare fiecăror firme înregistrate pe parcursul celor 4 ani, curățenia a constat doar în modificarea tipurilor parametrilor *CUI* și *CAEN* din tip numeric în tip string, după cum se poate observa și în figura de mai jos.

Pasul 2, reprezentat de combinarea datelor a avut loc în moduri diferite, astfel toate datele reprezentate de înregistrările indicatorilor economici ai firmelor pe parcursul celor 4 ani au fost supuse unui *UNION* în urma căruia am obținut un singur set de date cu toate înregistrările. Cât legat de datele înregistrate pe *Oficiul Național al Registrului Comerțului*, în urma curățării lor acestea au fost supuse aceleiași proceduri de *UNION* astfel încât s-au obținut toate datele necesare combinării atât a indicatorilor economici cât ți a denumirilor firmelor pe baza *CUI*-ului.

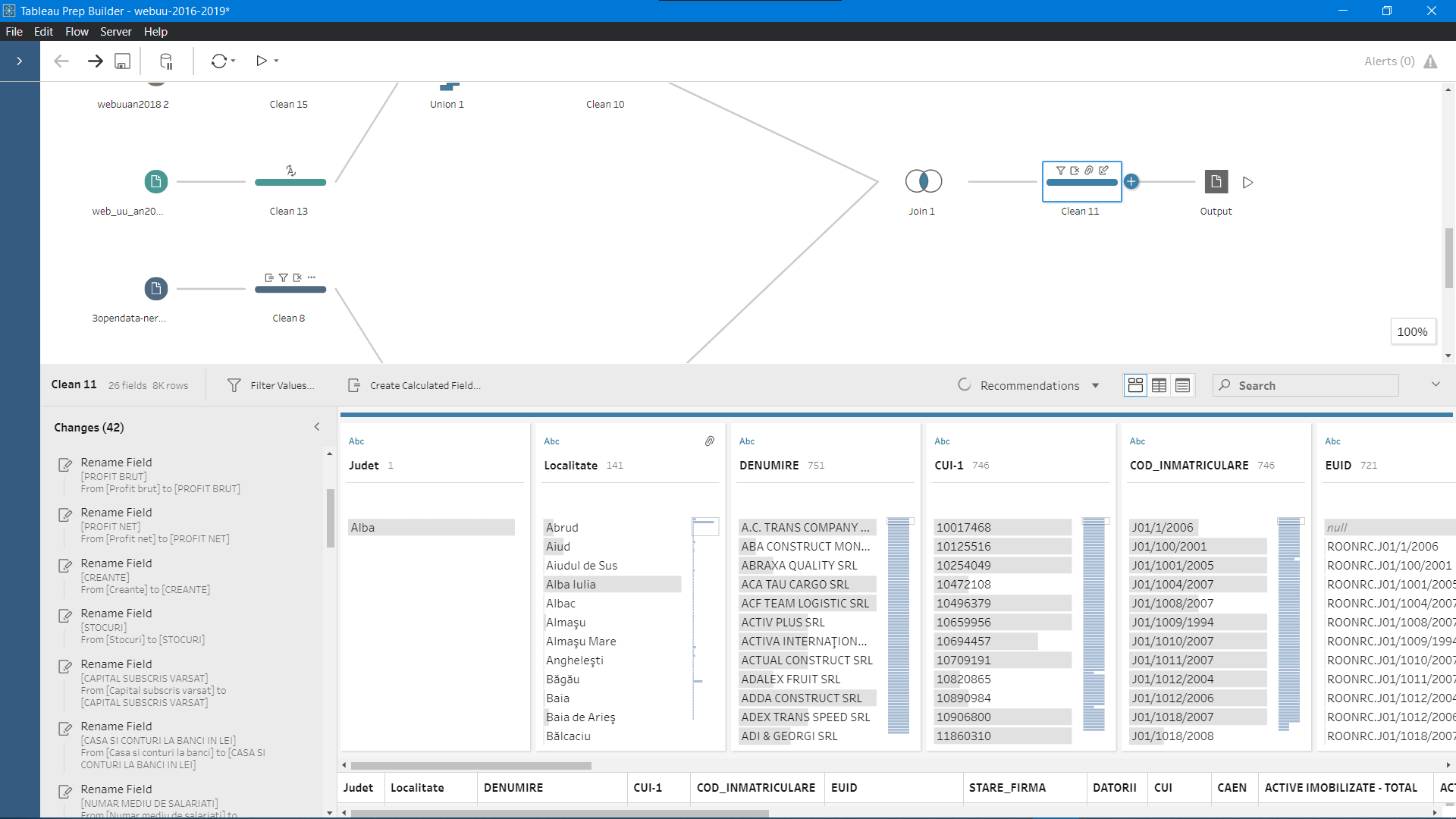


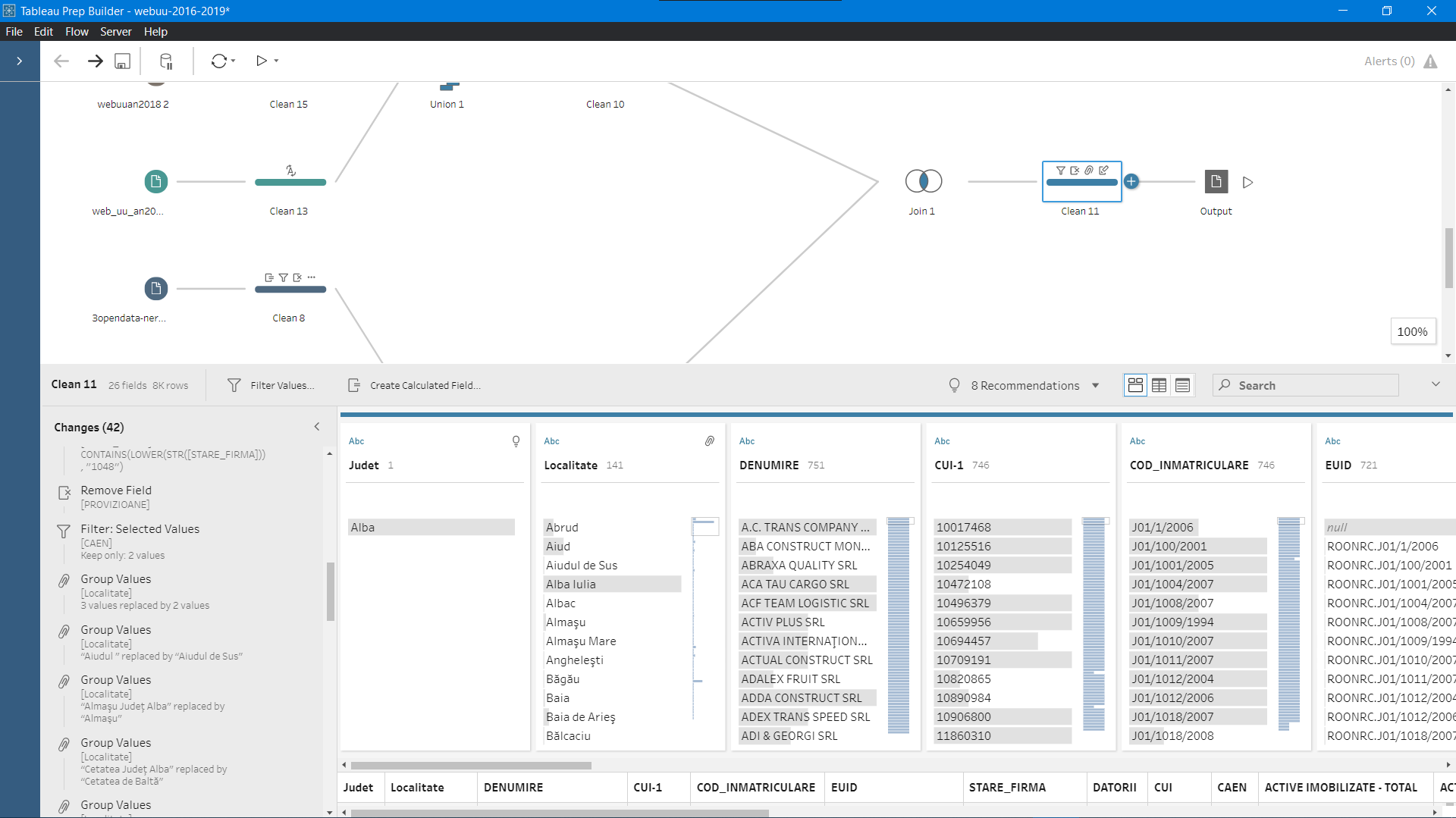
Nu în cele din urmă, aceste 2 fișiere rezultate au fost supuse procedurii de *JOIN*, în urma căruia au generat versiunea finală a setului de date. Tot acest proces poate fi observat în cadrul figurii de mai jos.



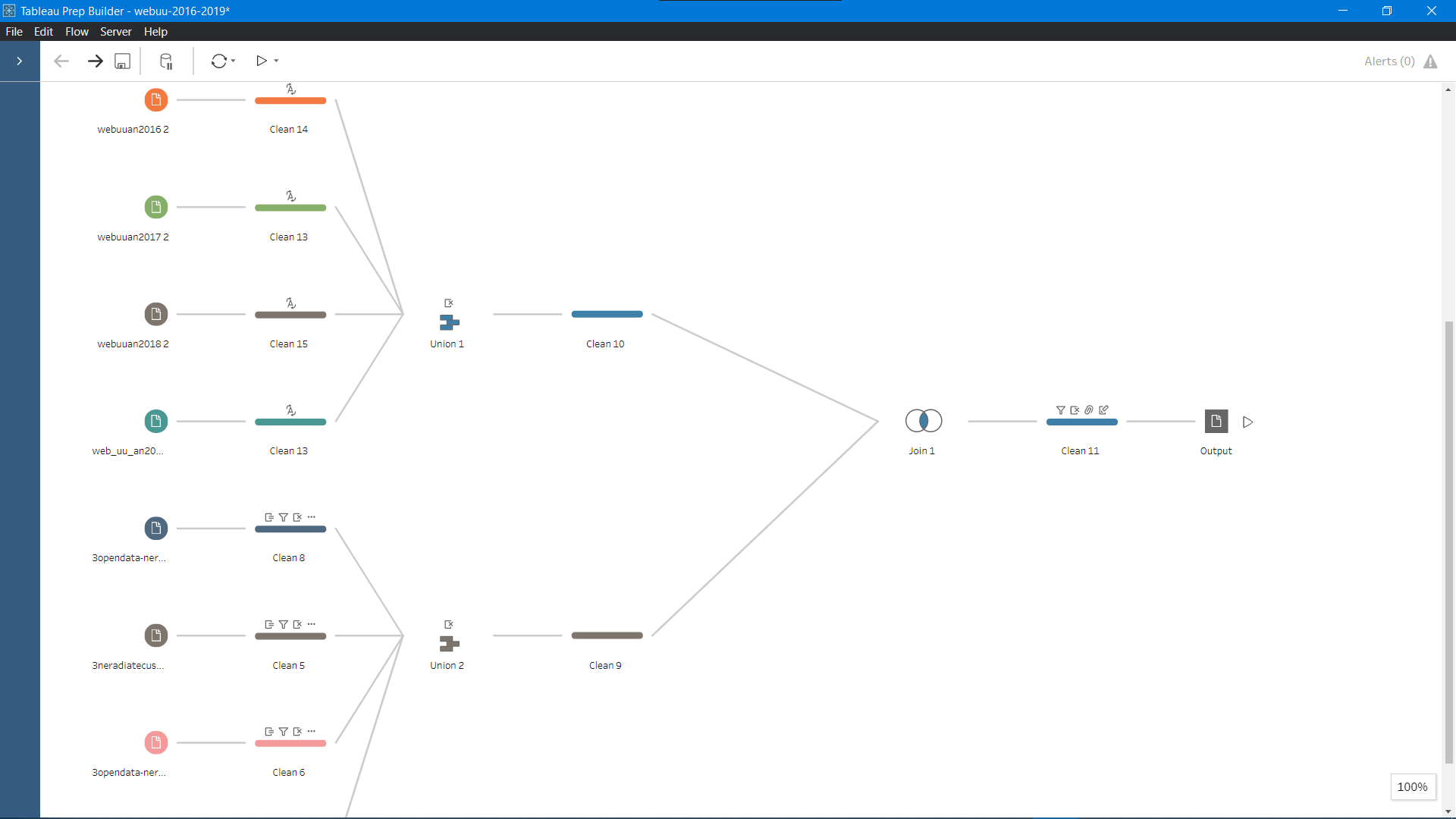
S-au obținut aproximativ 70 de mii de înregistrări care au necesitat din nou o curățare. Aceasta a constat în crearea de grupuri pentru localitățile care în urma *JOIN*-ului nu au fost identificate corect, ștergea câmpurilor a căror valori erau preponderent zero deoarece acestea ar nu ar fi generat nimic relevant în cadrul setului nostru de date și modificarea tipurilor de date ale unor înregistrări.

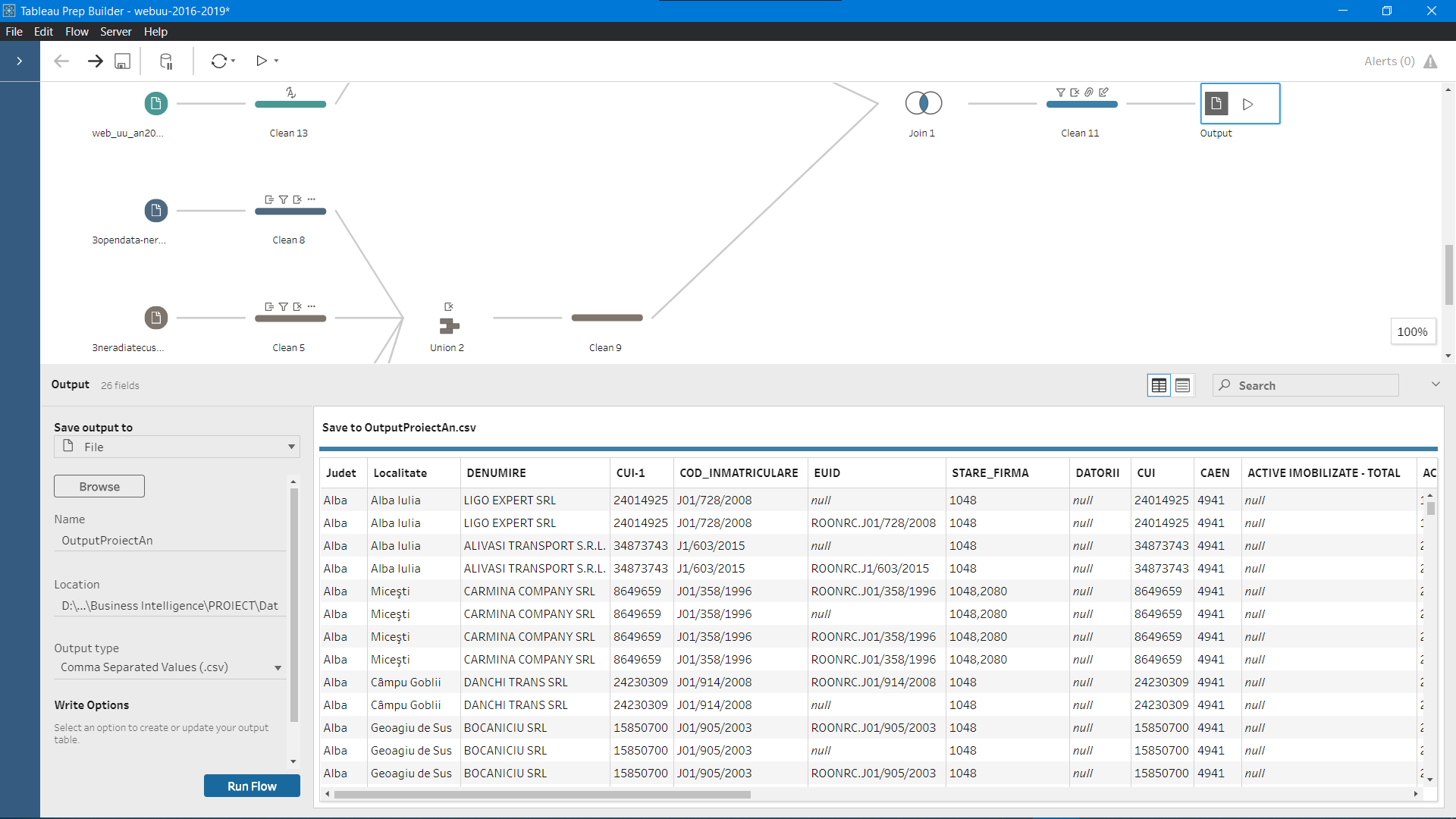
De asemenea, deoarece piața studiată cuprinde doar două coduri *CAEN*, s-a executat o selecție a codurilor *CAEN* din document, păstrând astfel dor cele două coduri relevante după cum se poate observa și mai jos.





Rezultatul tuturor acestor modificări a fost ulterior exportat sub forma unui output cu extensia csv iar mai apoi importat în Tableau Desktop pentru a urma procesul de analiză.



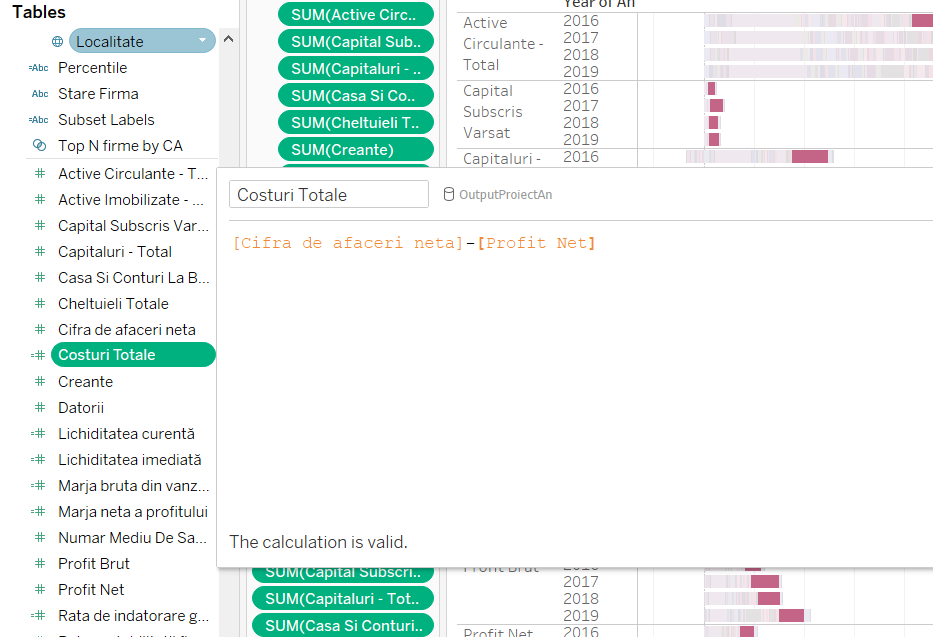


1. **ANALIZA DATELOR**

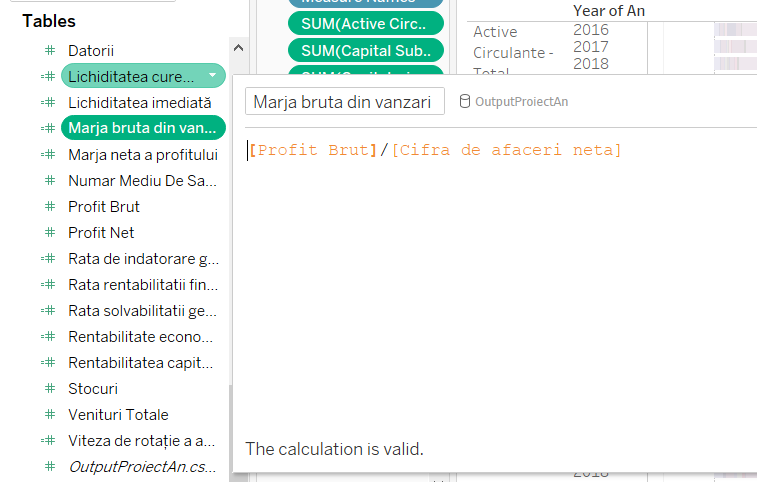
În urma conectării la setul nostru de date exportat din Tableau Prep, se dorește să se analizeze în detaliu o serie de probleme de business care ne ajută să înțelegem mai bine evoluția pe parcursul celor 4 ani a acestor firme, cât și să vedem care ar fi cel mai profitabil domeniu de activitate în județul Alba.

Pentru îndeplinirea cerințelor ce țin de lucrul cu foi, s-au calculat o serie de câmpuri, folosite ulterior în analiza și reprezentarea datelor.

* **Costul Total** – calculat ca și diferență între Cifra de afaceri netă și profitul net

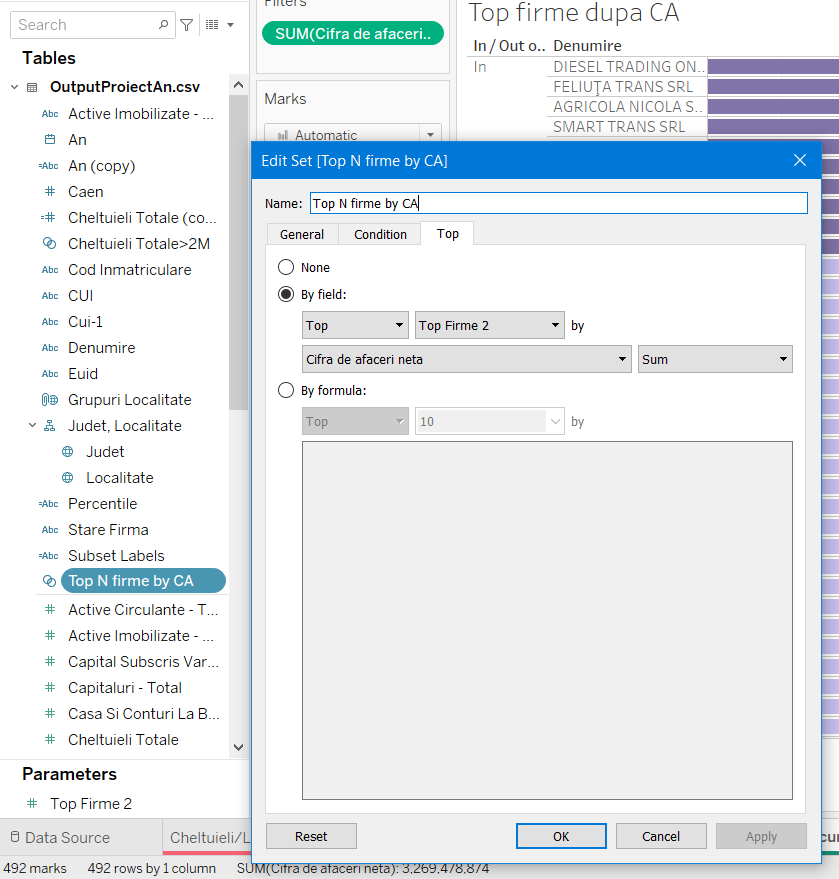


* **Marja brută din vânzări**, calculată ca și Profit Brut/Cifra de afaceri Netă



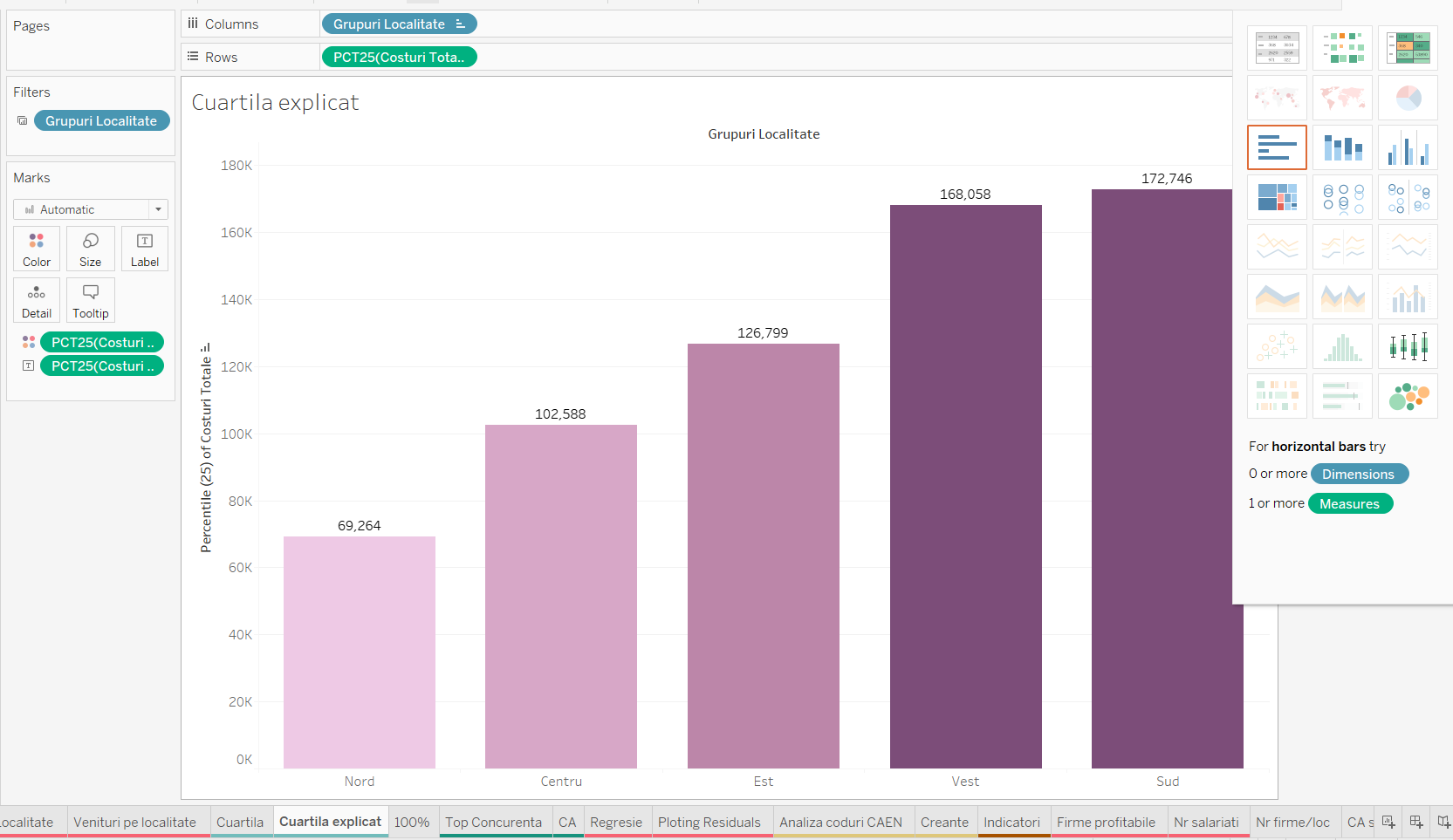
Parametrii:

* Parametru Top Firme care sortează firmele după cifra de afaceri, dinamic. Astfel putem vedea topul celor mai productive firme.

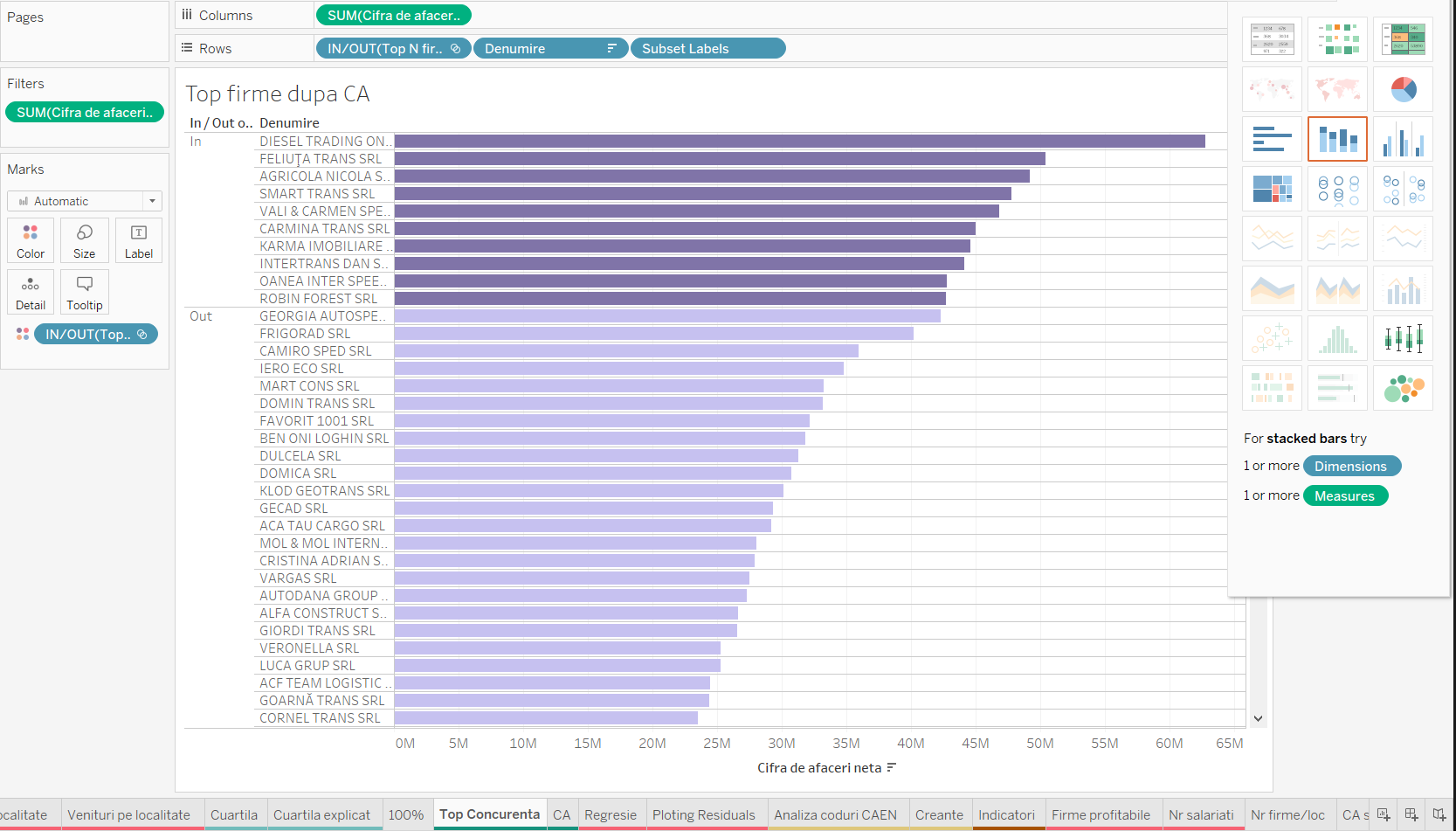


Tipuri de grafice folosite:

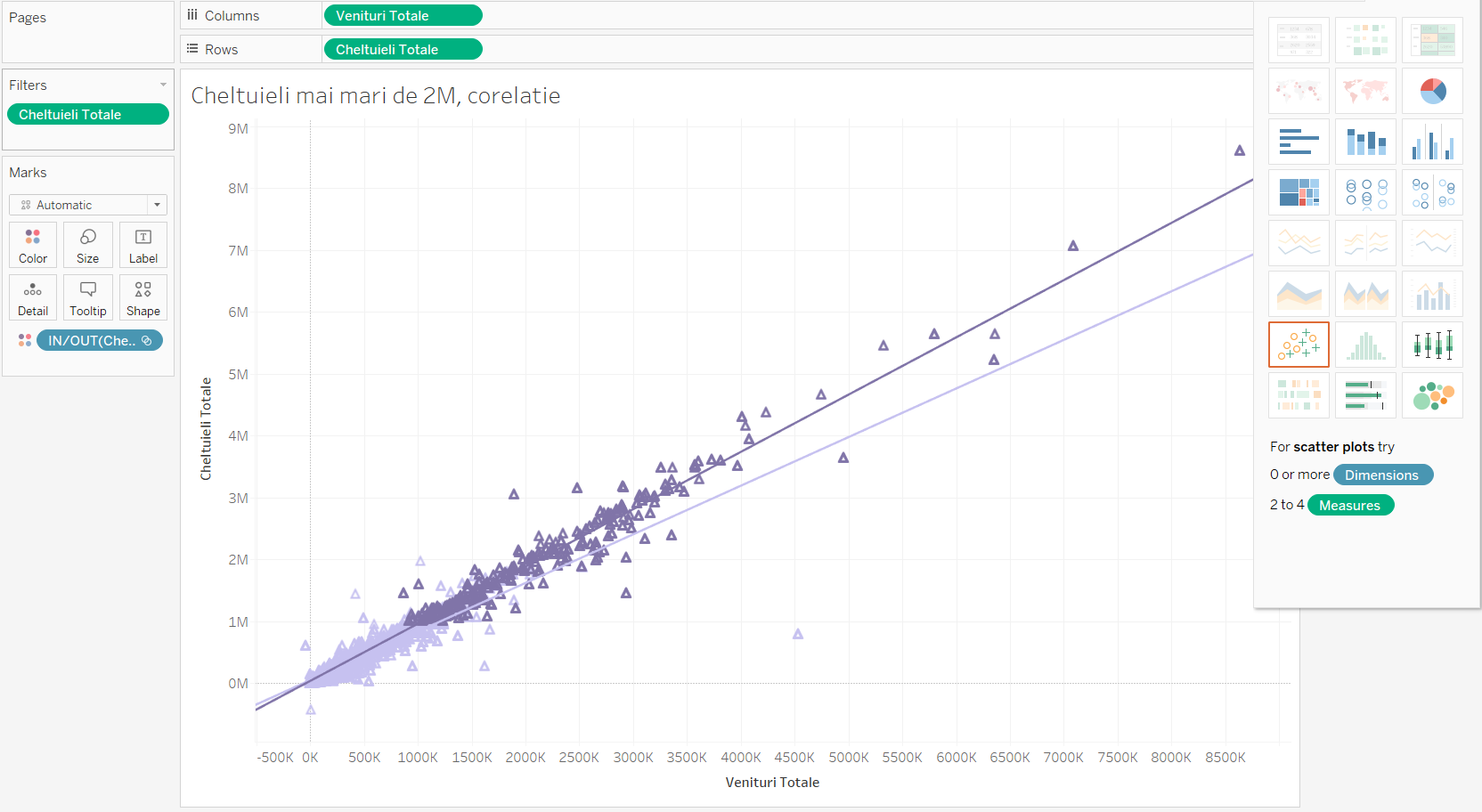
1. **Horizontal Bars**



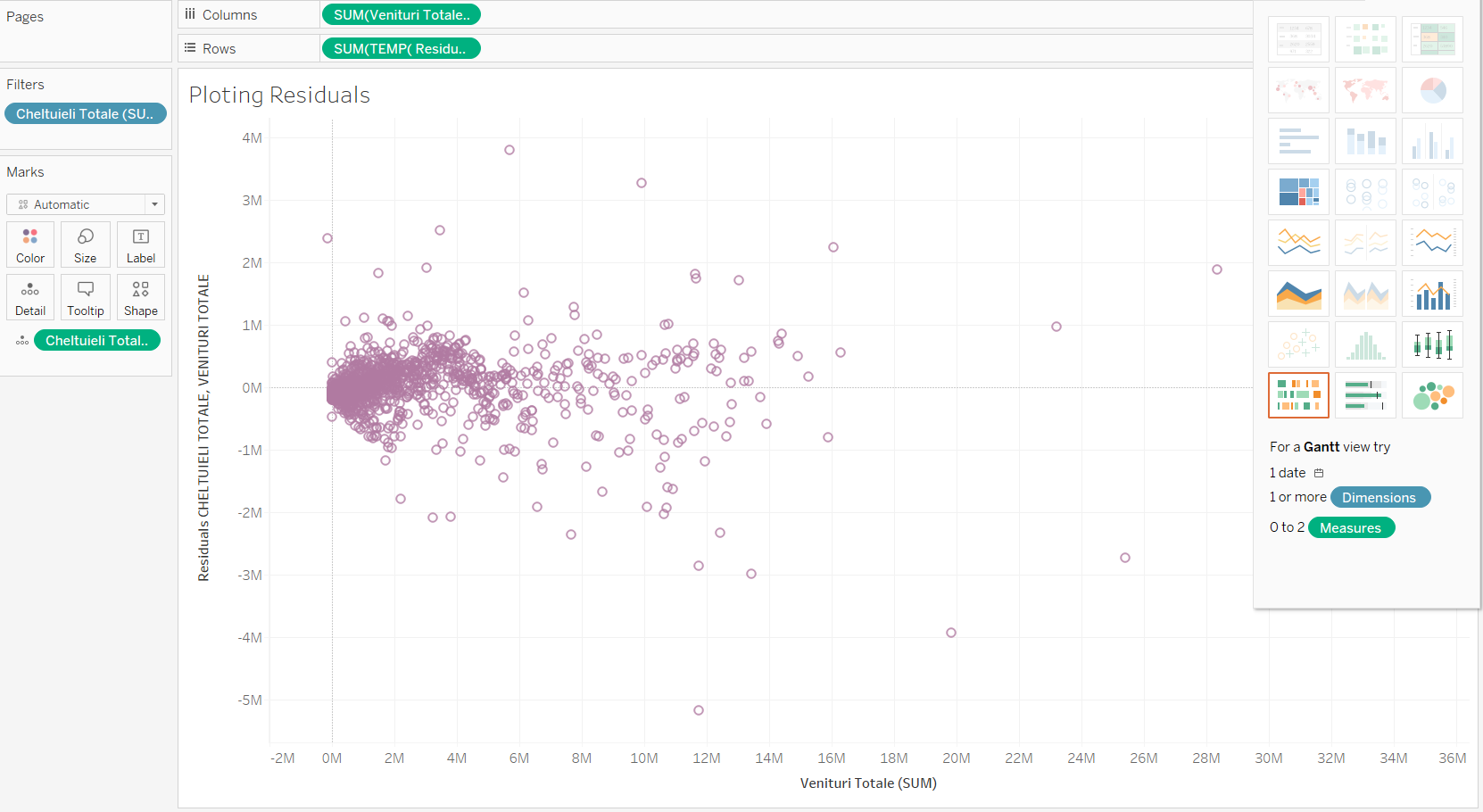
1. **Stacked Bars**



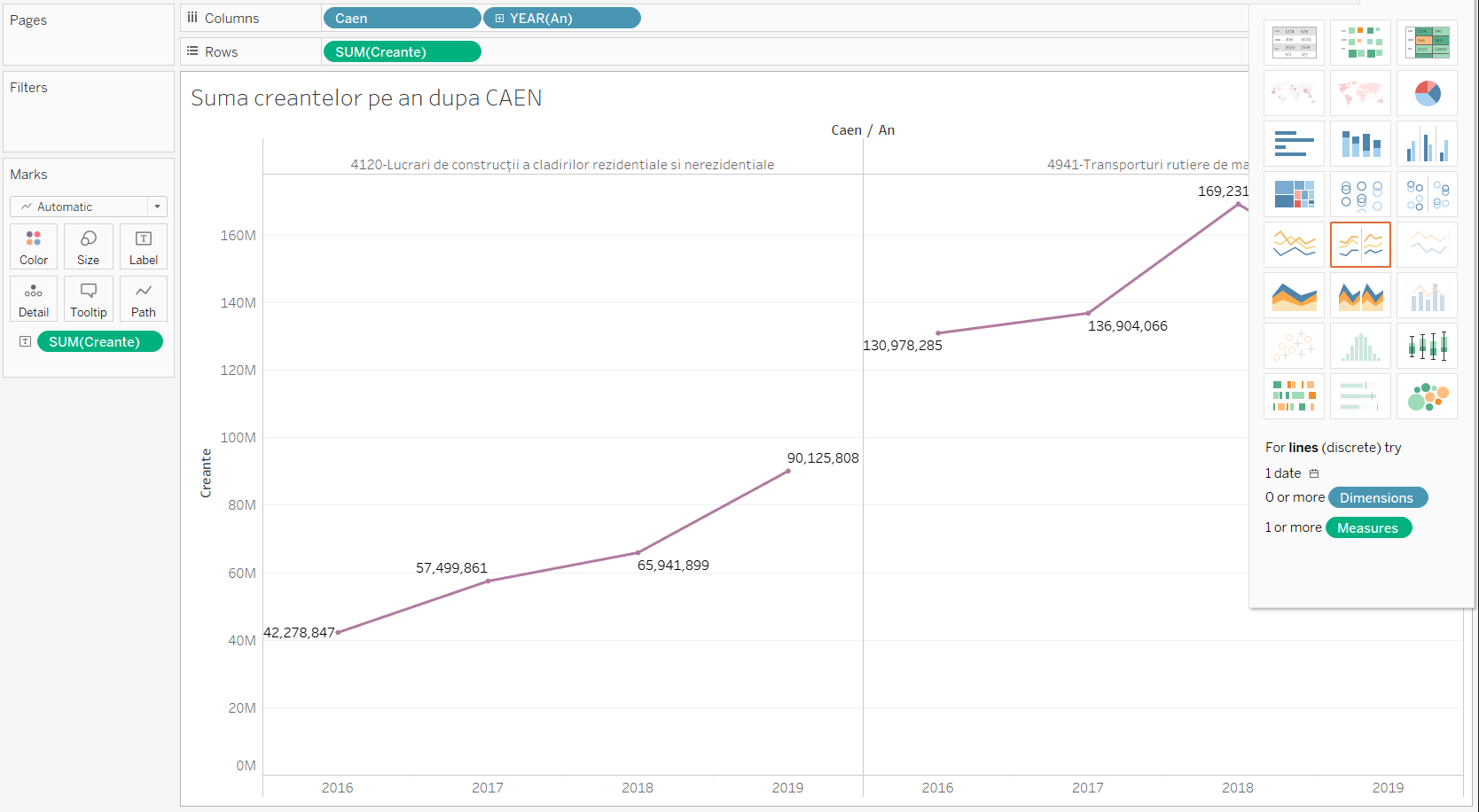
1. **Scatter Plots**



1. **Gantt**

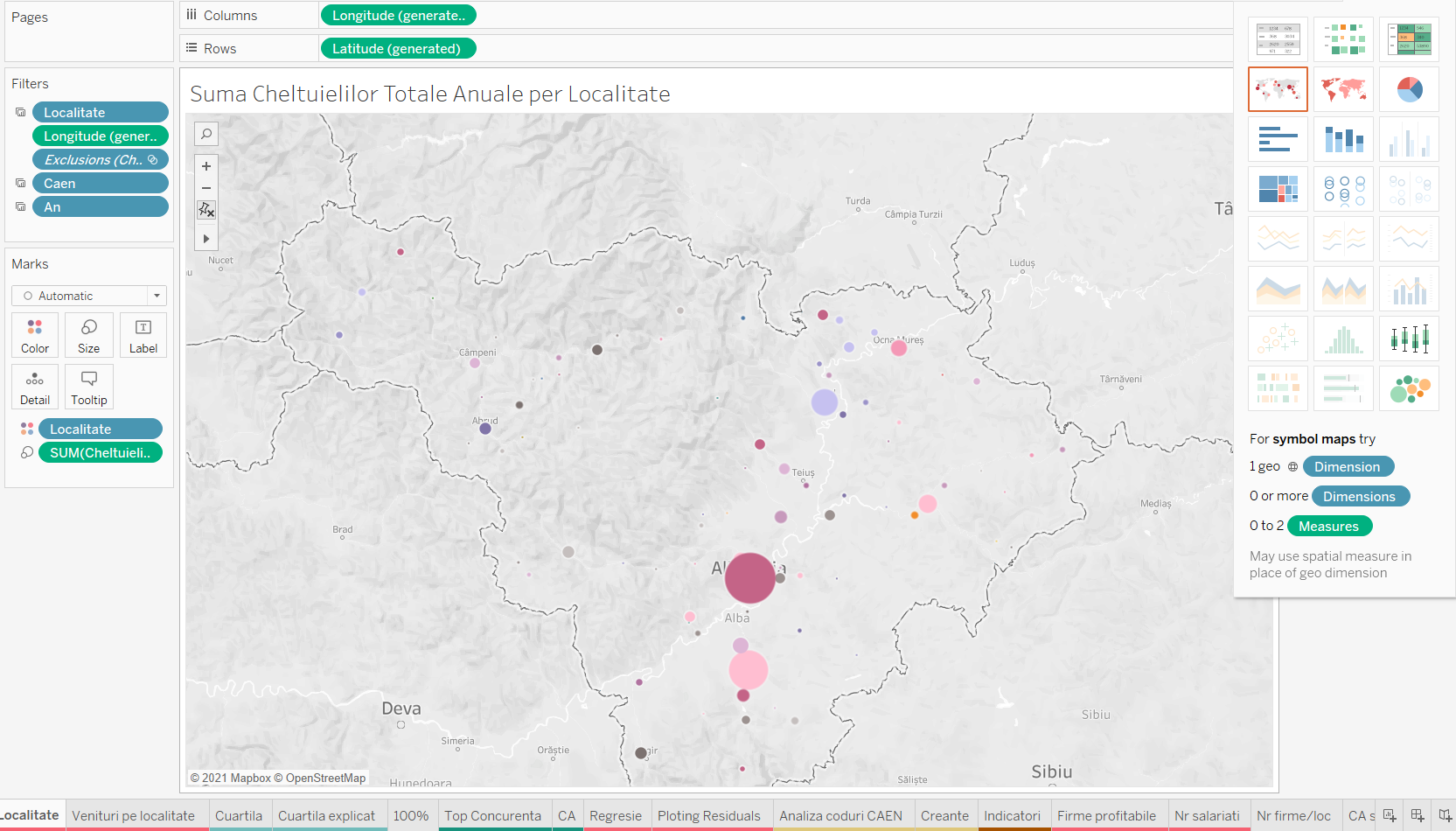


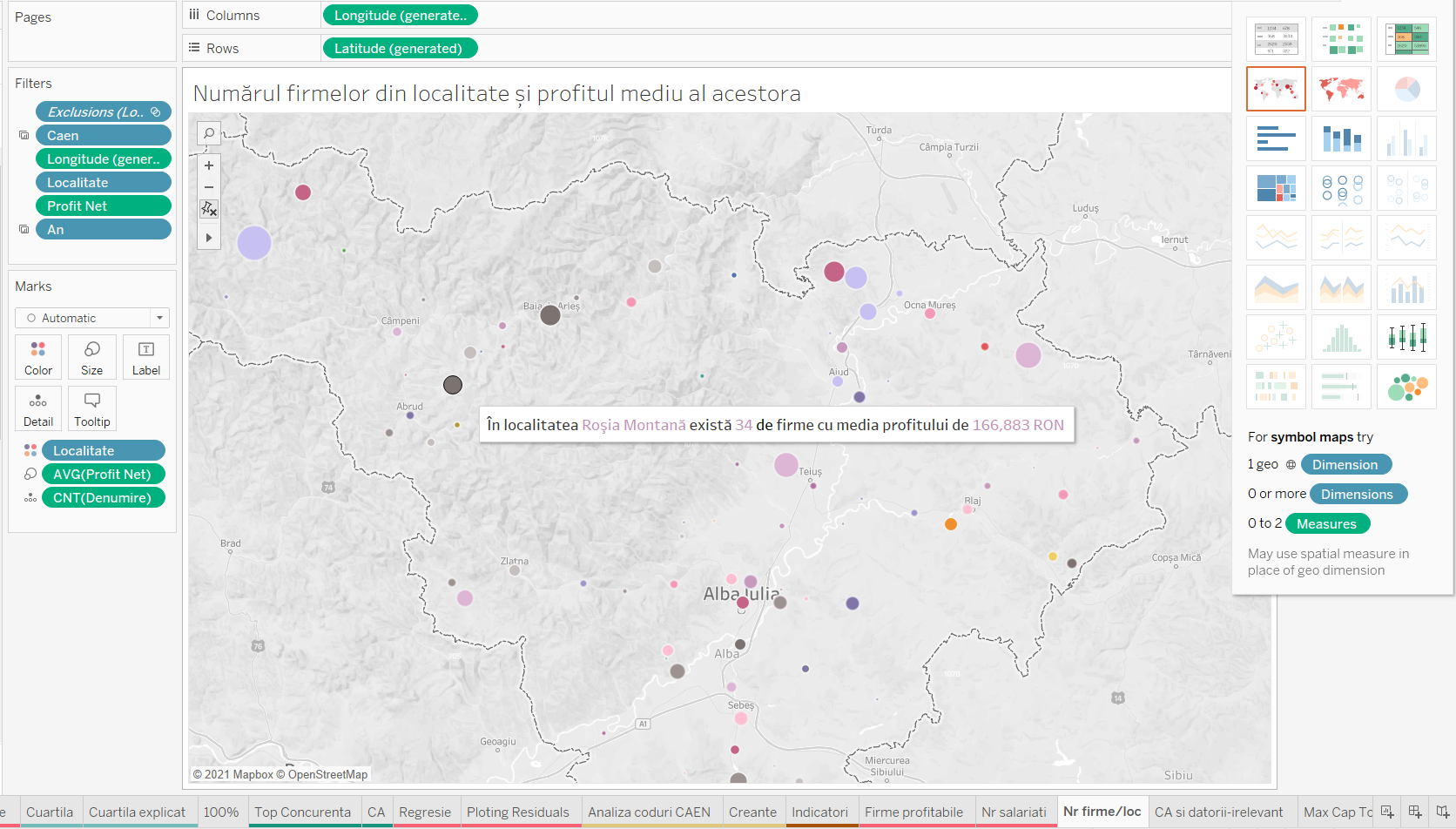
1. **Lines**



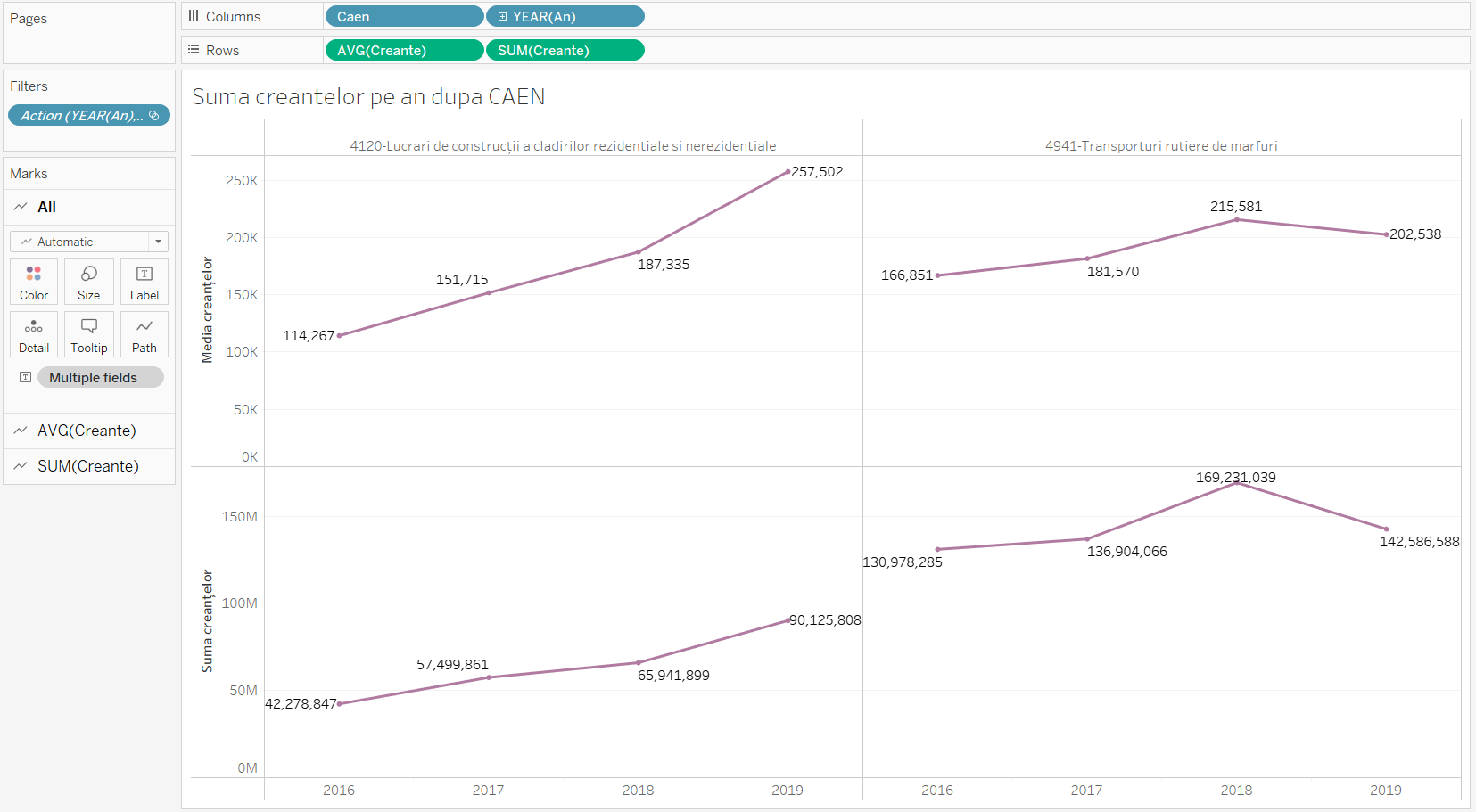
Hărți:

1. **Symbol Maps**





Analiza de trend pentru evoluția în timp a datelor:



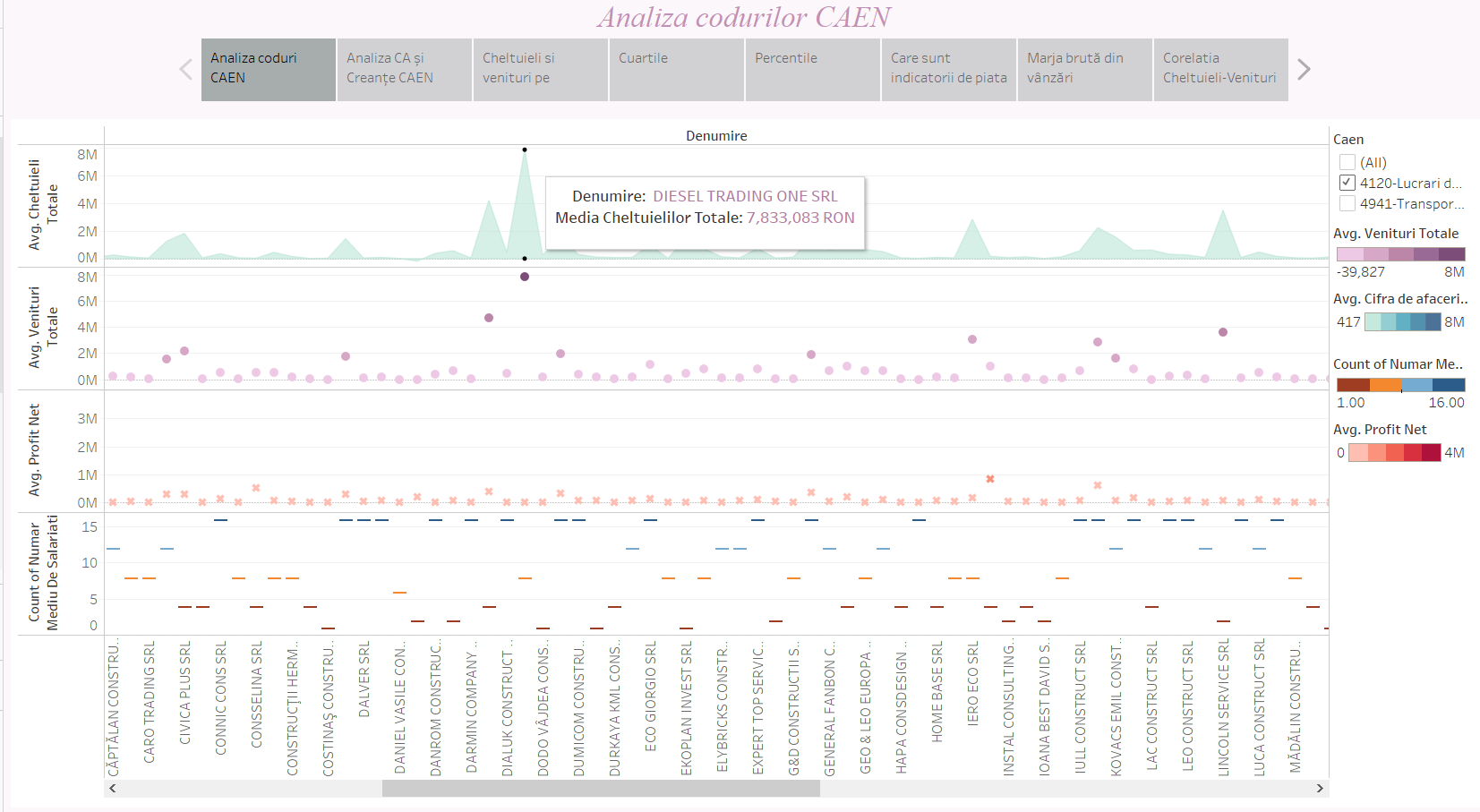
Astfel, ca și pas următor se adresează o serie de întrebări de business la care vom răspunde pe parcursul acestui document cu ajutorul Dashboard-urilor și a Story-urilor.

* Care dintre cel 2 coduri CAEN este mai relevant pe județul selectat ?
* Care este cel mai profitabil CAEN în cei 4 ani ?
* Care este nivelul creanțelor firmelor pe parcursul celor 4 ani ?
* Care sunt veniturile, respectiv cheltuielile pe o localitate pe an în funcție de CAEN ?
* Cât la % din costurile totale din zonele județului sunt prezente în prima quartilă ?
* Care este numărul de firme din top 25%, middle 25% și bottom 25% ?
* Care sunt indicatorii de piață cei mai relevanți în cazul firmelor dintr-un anumit județ pe an ?
* Există o corelație între venituri și cifra de afaceri pe datele înregistrate ? Dacă da, ce tip de corelație ?
* Care sunt top 10 companii concurente în cadrul CAEN-ului cel mai profitabil ?
* Câte firme sunt și ce profit au în cadrul fiecărui oraș/localitate după CAEN ?
* Concluzie: care este cel mai bun domeniu de activitate dintre cele 2 CAEN-uri pentru a deschide o firmă în anul 2021 ?

1. **Care dintre cel 2 coduri CAEN este mai relevant pe județul selectat ?**

Pentru a putea răspunde la această întrebare am recurs la analiza codurilor CAEN și a celor mai relevanți indicatori ai pieței. S-a decis acordarea atenției asupra datoriilor, veniturilor totale, profitului net cât și numărului mediu de angajați pentru cele 2 CAEN-uri. Astfel, s-a constatat faptul că veniturile mari atrag după sine cheltuieli pe măsură, profitul net este direct proporțional cu venitul, însă nu este suficient. O analiză mai exactă constă în compararea exactă pe două exemple:

* ALFA CONSTRUCT SRL deține un nivel mediu al veniturilor totale de 2 219 519 RON, în timp ce acestea atrag după sine datorii de doar 557 007 RON, însă media profitului net este de 63 902 RON, ceea ce reprezintă doar 2,87% din valoarea medie a veniturilor obținute de această firmă.
* Surpriza apare la nivelul altei firme, DIESEL TRADING ONE SRL care deține un nivel mediu al veniturilor totale de 7 856 946 RON pe când nivelul profitului net este de doar 18 658 RON, doar 0.23% din media profitului, o valoare extrem de mică, ceea ce ne duce cu gândul la o valoare mare în rândul cheltuielilor, iar într-adevăr se poate observa că nivelul mediu al cheltuielilor este de 7 833 083 RON.

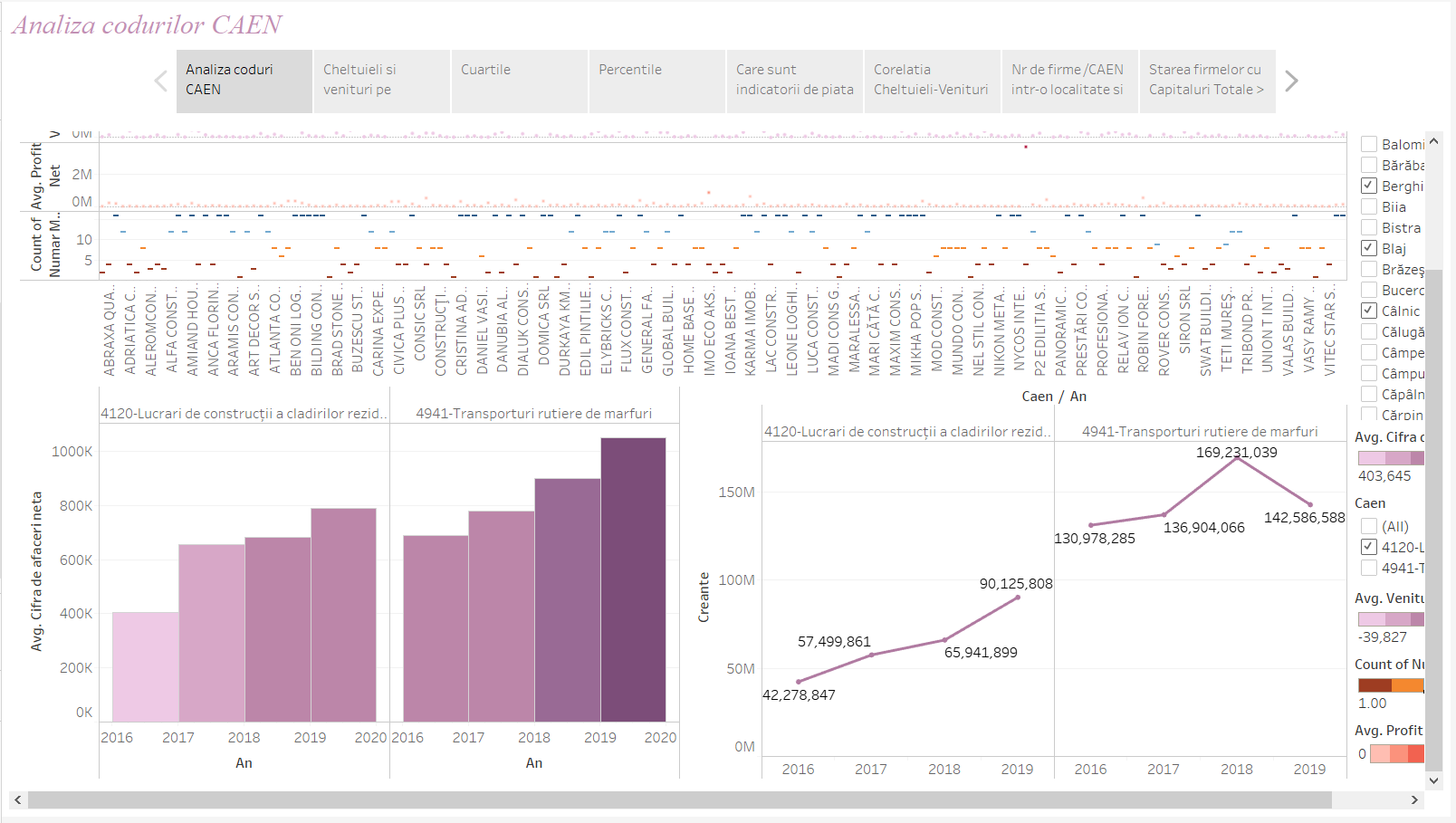


1. **Care este cel mai profitabil CAEN în cei 4 ani ?**

Pentru această întrebare s-au luat în calcul la fel ca și la pasul anterior repartizarea chltuielilor, veniturilor, creanțelor, a profitului și a mediei cifrei de afaceri pe ani, informații care adunat și puse cap la cap duc la concluzia că domeniul de activita Transporturi rutiere de mărfuri este mult mai profitabil, atât din punct de vedere al valorii datoriilor, a profitului net cât și al cifrei de afaceri care deși a suferit o scădere dramatică în decursul anului 2018-2019, este de 1.58 ori mai mare decât cifra de afaceri corespunzătoare domeniului concurent.

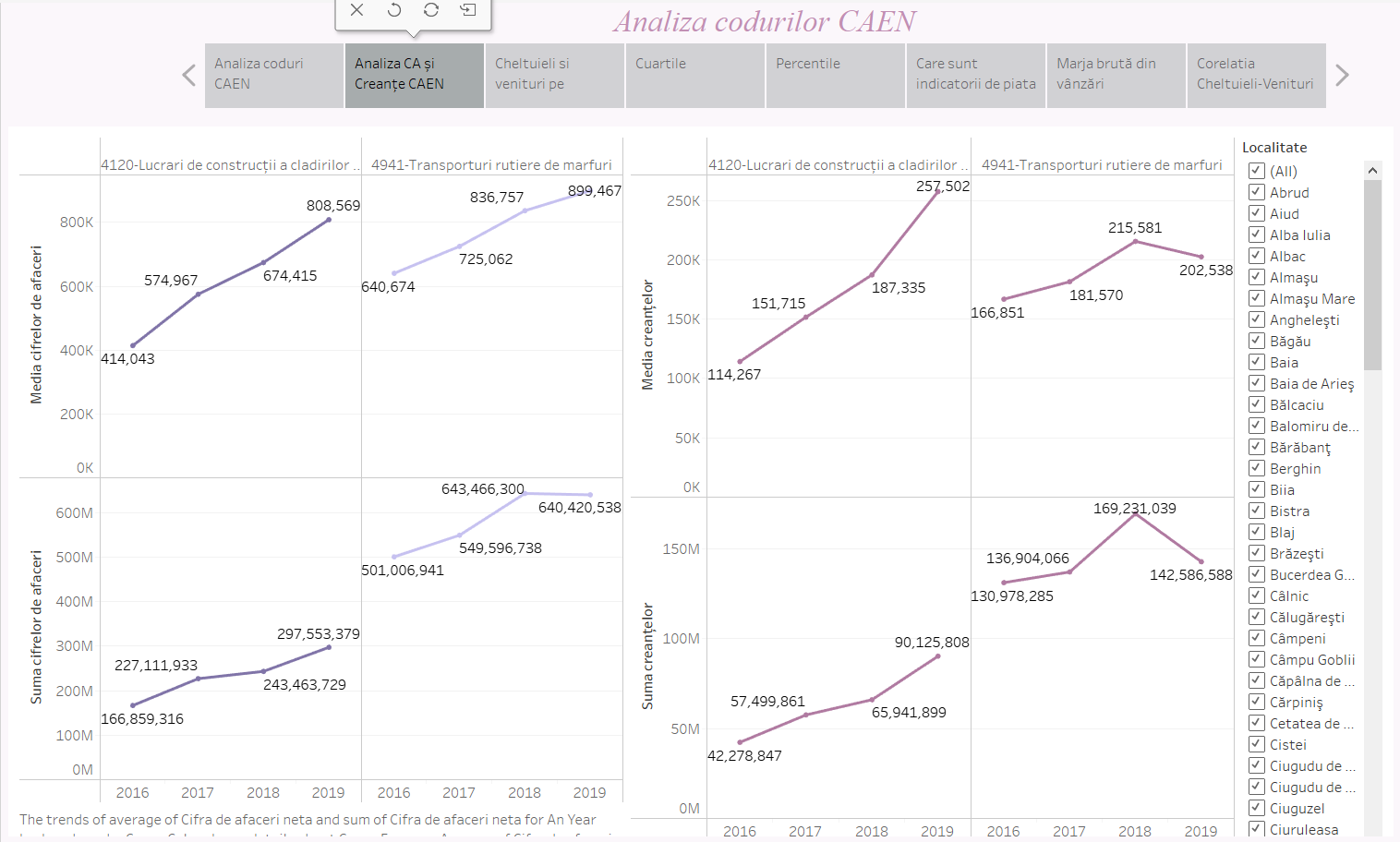
* Pentru CAEN 4120 – Lucrări de construcții a clădirilor rezidențiale, creșterea a fost aproape liniară de la o valoare inițială a creanțelor în anul 2016 de 42,278,874 RON la 90,125,808 RON, o creștere de aproximativ 113.16%
* Pentru CAEN 4941 – Transporturi rutiere de mărfuri însă, creșterea a fost mai mare pe parcursul anului 2017-2018, o creștere de 23.61% față de anul imediat următor în care aceasta a scăzut cu aproximativ 15.74%

Conform mediei cifrei de afaceri, din acest punct de vedere cel mai rerezentativ CAEN este 4941-Tranporturi rutiere de mărfuri pe anul 2019, unde media cifrei de afaceri este de 1 050 862 RON, comparativ cu CAEN-ul 4120-Lucrări de construcții a clădirilor rezidențiale unde media cifrei de afaceri pe anul 2019 este de doar 788 337 RON.



1. **Care este nivelul creanțelor firmelor pe parcursul celor 4 ani ?**

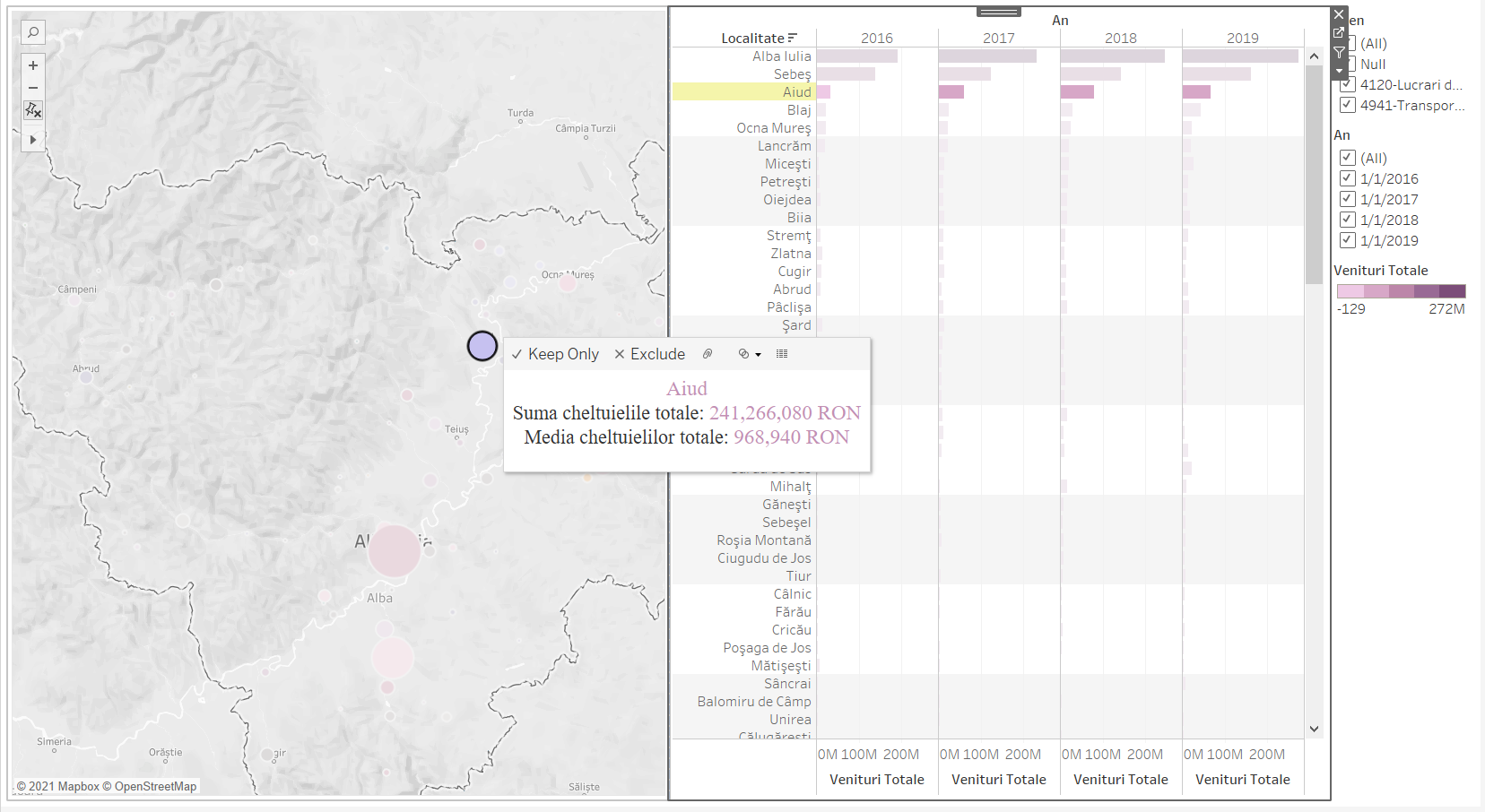
Această întrebare este analizată în cadrul aceluiași dashboard, iar după cum s-a modificat și cifra de afaceri putem observa că aceasta a avut un impact direct asupra creanțelor, în sensul în care ele urmează același curs ca și creanțele.



Astfel, s-a dorit compararea atât a mediei cifrelor de afaceri dintr-o localitate cât și a creanțelor, pentru a putea structura o imagine de ansamblu asupra procesului de conversie a creanțelor în capital social. În urma acestui proces, angajații primesc o parte echitabilă din acțiunile firmei în schimbul aportului în bunuri sau bani în capitalul social.

1. **Care sunt veniturile, respectiv cheltuielile pe o localitate într-un an după CAEN ?**

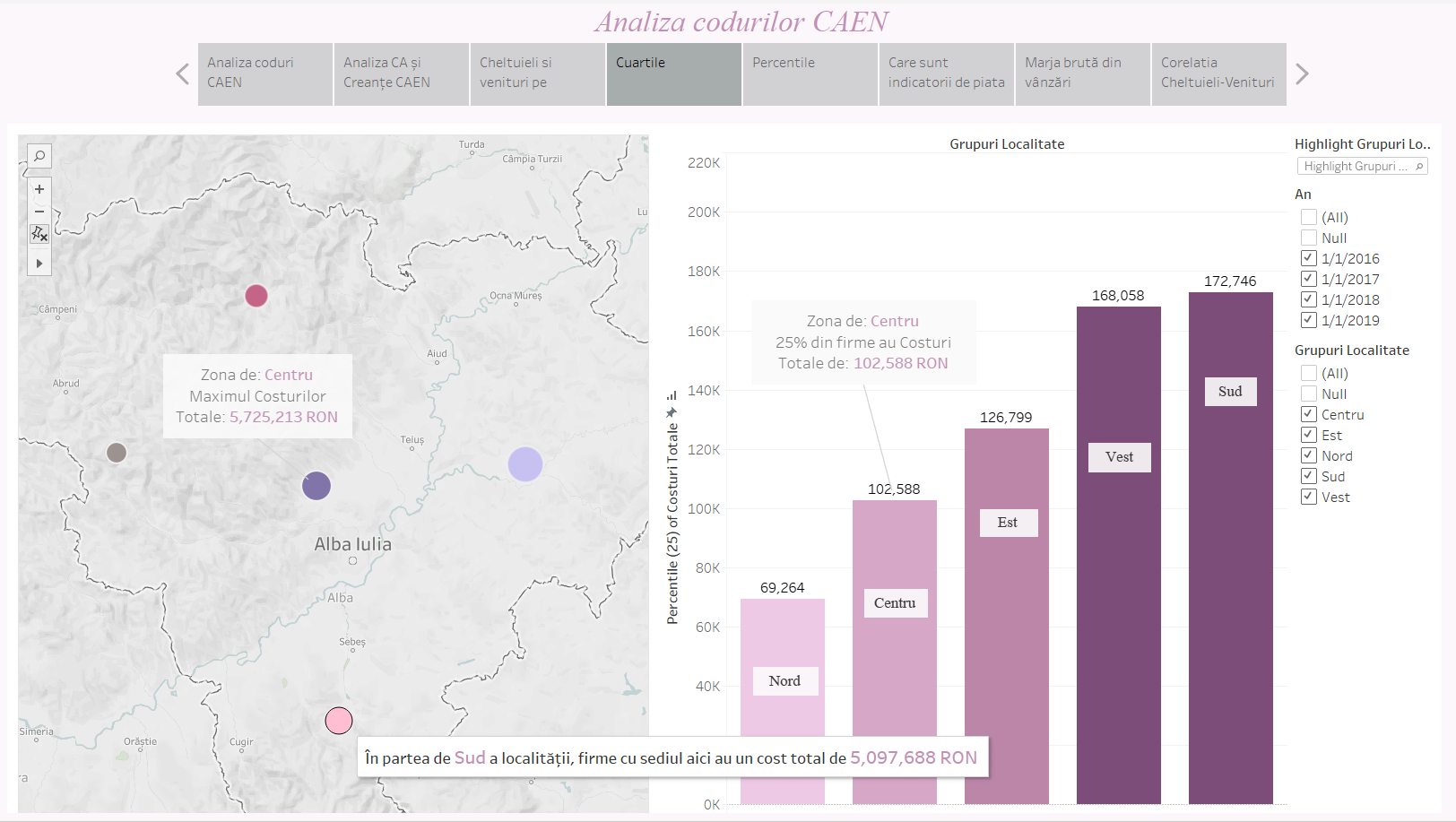
Pentru a răspunde la acestă întrebare ne-am folosit de indicatorul veniturilor totale, respectiv al cheltuielilor totale în cadrul unei hărți pentru a ne creea o imagine grafică a acestor indicatori. Astfel aceste date pot fi observate cu ușurință în cadrul dashboard-ului.



Conform graficului, pentru localitatea Aiud s-au înregistrat valorile cumulate a cheltuielilor totale de 241 266 080 RON cu o medie a acestora de 968 940 RON pe toți cei 4 ani, cât și valoarea cumulată și medie a veniturilor totale pe parcursul celor 4 ani.

1. **Cât la % din costurile totale din zonele județului sunt prezente în prima quartilă ?**

În cadrul acestei întrebări ne-am folosit de indicatorul costurilor totale, calculat ca și Cifra de afaceri netă-Profit Net. Acesta a fost alocat pe grafic celor 4 ani cu ajutorul filtrului de localități astfel încât se poate observa că cel mai profitabil oraș este Alba Iulia, cu suma veniturilor totale pe anul 2019 de 272 396 615 RON.

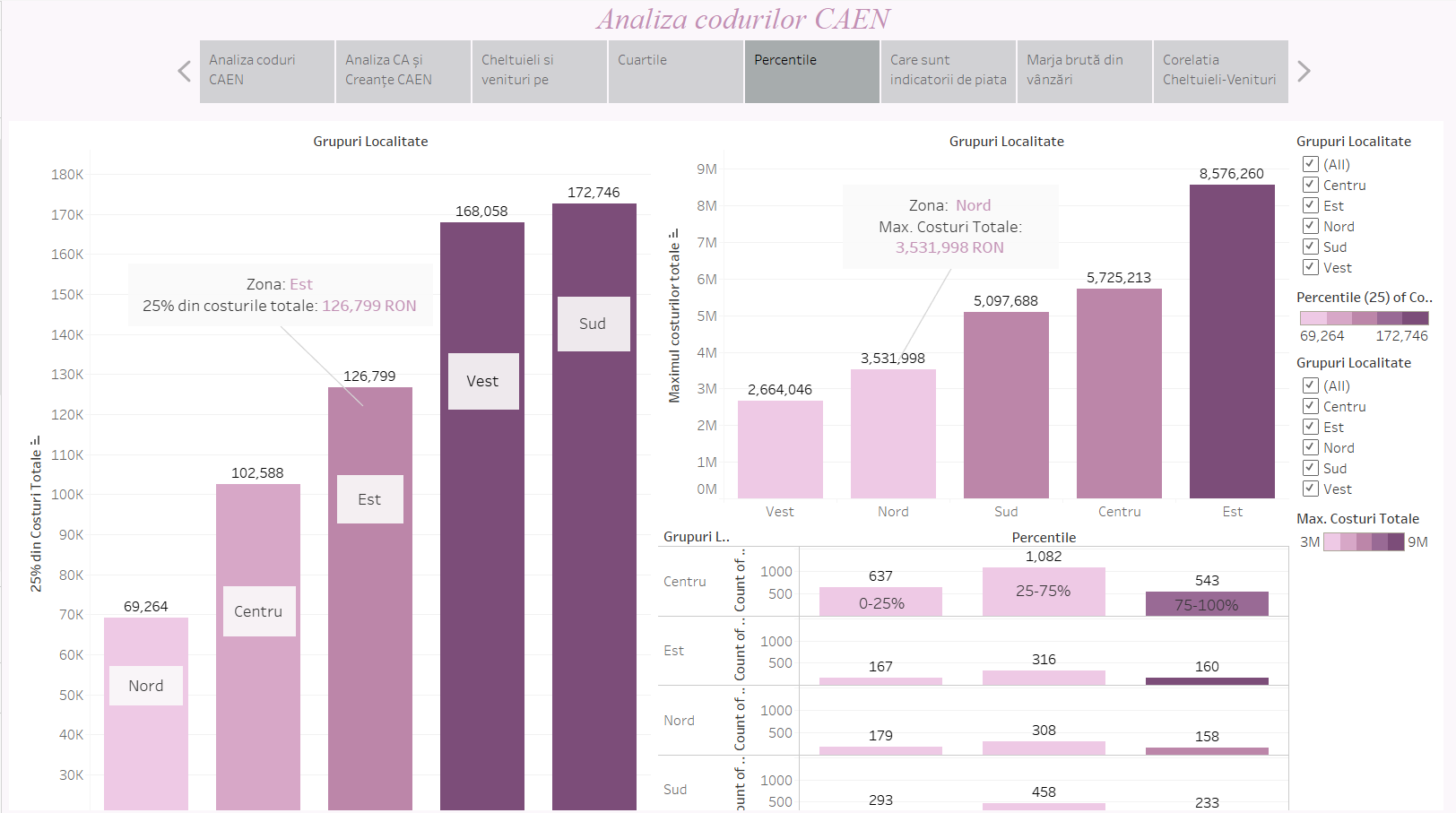


Calculând procentajul reprezentat din costurile totale din toate cele 4 cuartile, observăm că în Nord spre exemplu, acest cost de 69 264 RON reprezintă 1.96% din costul total pe acest sector, e 3 531 998 RON. În partea de Este, costul total maxim este de 8 576 260 RON, iar 25% din firme prezintă un cost de 126 799 RON, adică de 1.47 ori mai mic.

În zona din Centru, avem un maxim al costurilor totale înregistrat de 5 725 213 RON iar 25% din firme dețin o valoare de doar 1.79% din acest cost.

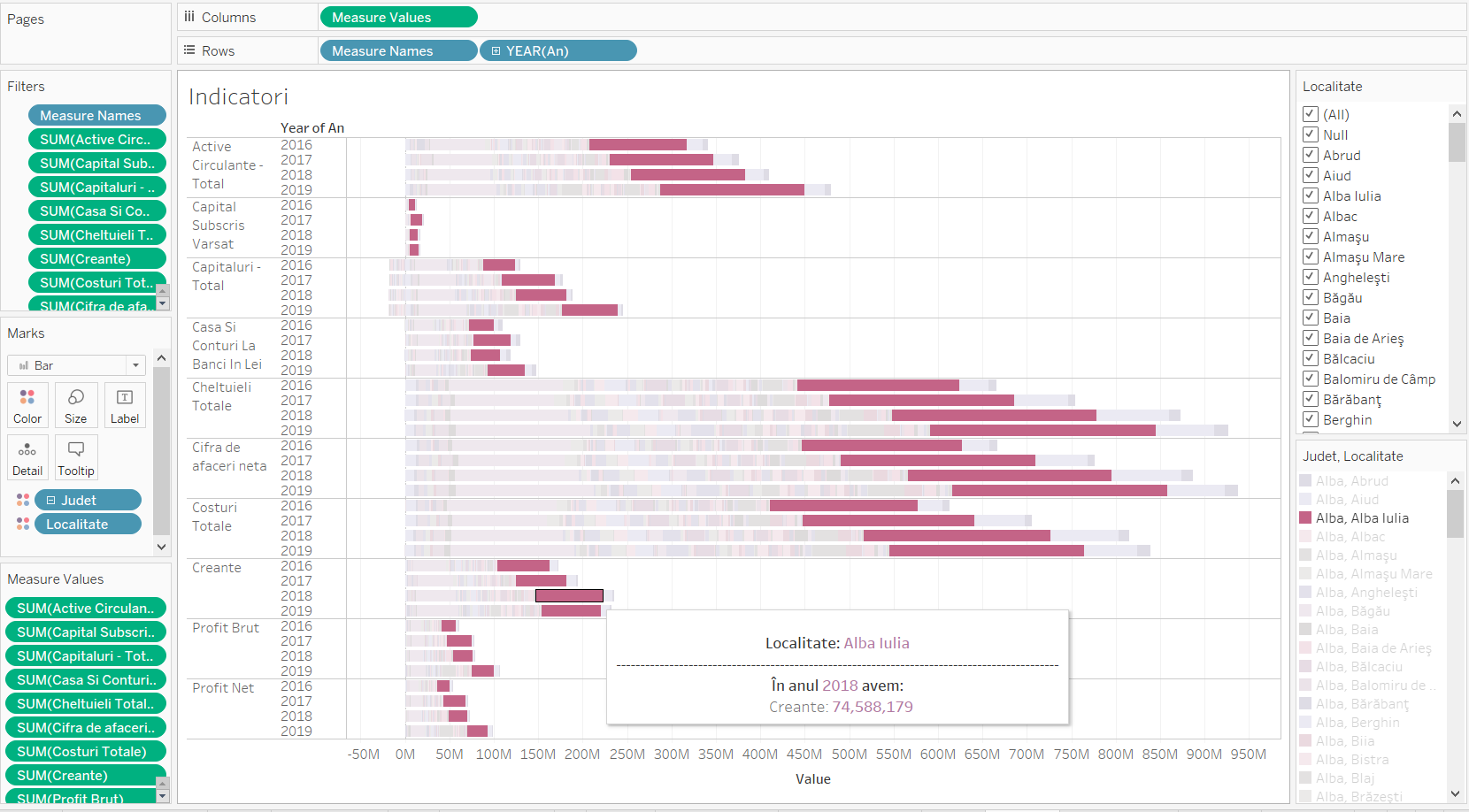
1. **Care este numărul de firme din top 25%, middle 25% și bottom 25% ?**

Pentru a răspunde la această întrebarea a fost necesară alocarea localităților pe grupuri geografice, astfel încât împărțirea lor s-a efectuat astfel: Nord, Est, Sud, Vest și Centru. Pentru fiecare zonă geografică s-a conturat indicatorul economic al costurilor totale, din cadrul primei cuartile, cât și numărul de firme aflate în e percentile. Astfel, se poate deduce faptul că în partea din centru, 25% din firme se încadrează în partea de Bottom cuprinsă între 0-102 298,3 RON, 50% din firme se încadrează ân partea de Middle, cu o valoare a costurilor totale de 103 000-815 997RON, iar 25% din firme se încadrează în Top 25, adică au valoarea costurilor totale între 816 000-5 725 213 RON, ceea ce înseamnă că reprezentarea respectă o distribuție Gaussiană.

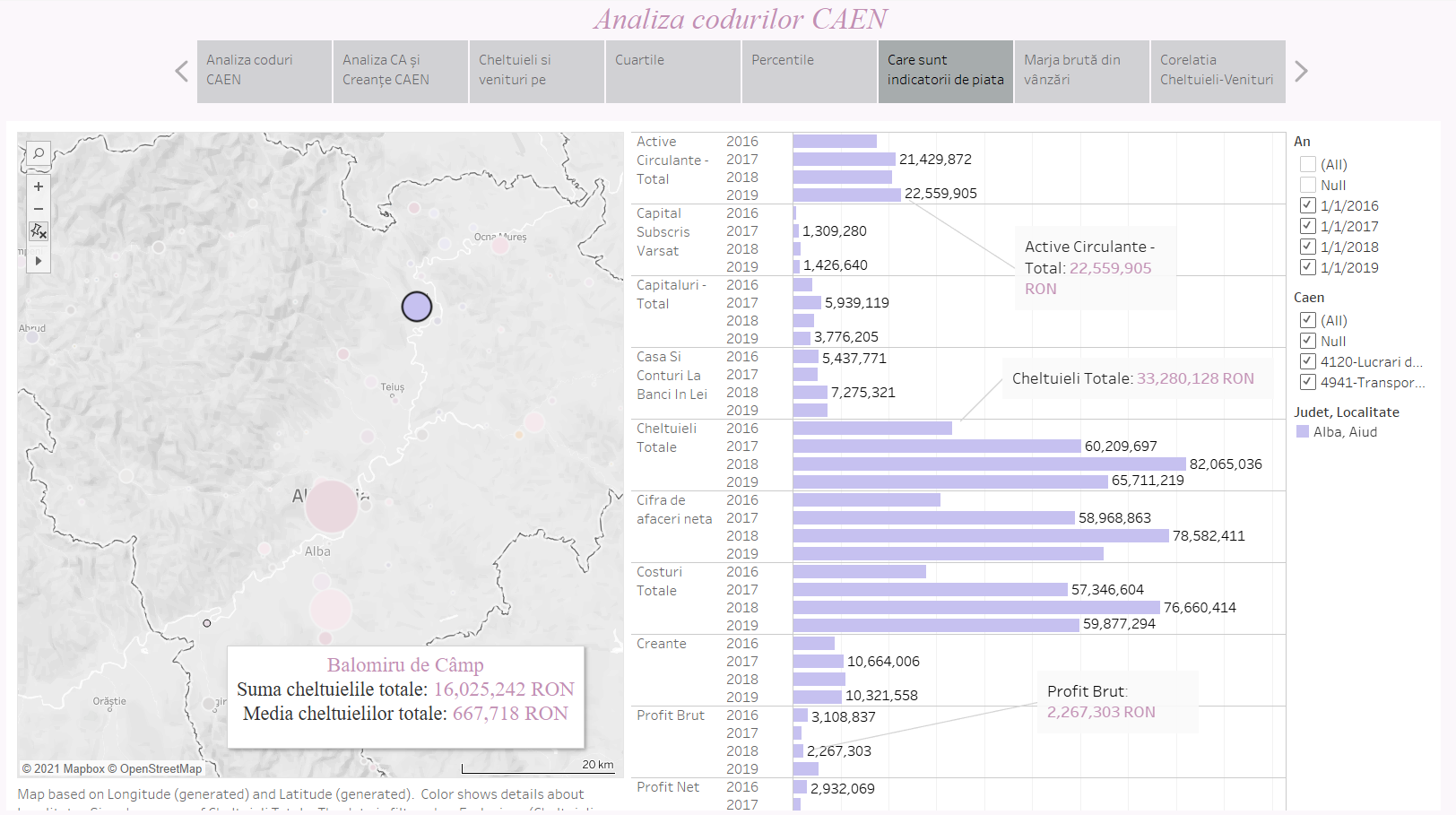


1. **Care sunt indicatorii de piață cei mai relevanți în cazul firmelor dintr-un anumit județ pe an ?**

Astfel încât există o serie de indicatori mai relevanți s-a decis alocarea lor în cadrul unui worksheet special, după mai multe filtre și marcatori, chiar dacă acesta va reprezenta în final un grafic mai complex, Legătura se va face printr-o acțiune la nivelul hărții care va selecta locația corespunzătoare în momentul redirectării pe acest worksheet, iar lucrurile vor fi mult mai clar vizibile astfel. S-a efectuat o analiză a indicatorilor: Active Circulante, Capital Subscris Vărsat, Capitaluri Total, Casa și Conturi la Bănci, Cheltuieli Totale, Cifra de Afaceri Netă, Costuri Totale, Creanțe și Profit Brut/Net. În cadrul graficului rezultat se pot observa valorile individuale ale acesto indicatori. Un exemplu concret ar fi Creanțele pe anul 2018 din localitatea Alba Iulia cu o valoare de 74 588 179 RON.

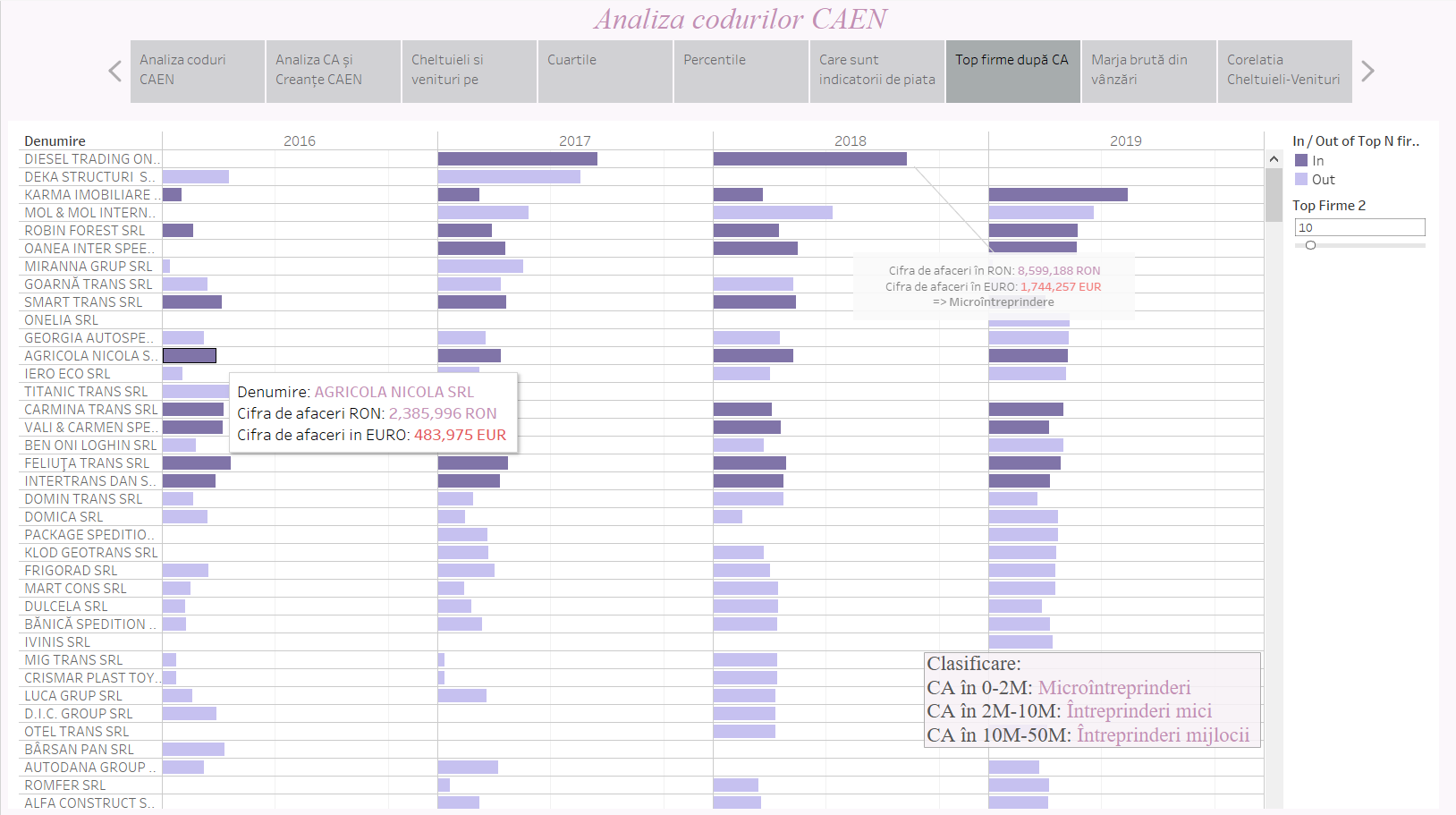


În cadrul story-ului se poate observa mult mai bine acest lucru deoarece selecția se face după localitate, astfel încât diferența devine vizibilă.



1. **Care sunt top 10 companii concurente în cadrul CAEN-ului cel mai profitabil și în ce categorie de întrprinderi se situează ?**

Topul celor mai profitabile firme după CA s-a realizat prin crearea unui parametru care selectează firmele cu o cifră de afaceri maximă și le ordonează descrescător pe fiecare an. Astfel se poate observa că firmele care se regăsesc în top în anul 2016 își păstrează poziția și pe parcursul anilor următori, însă nu avem în județ firme care să se situeze în clasa întreprinderilor mici, fapt pentru care toate acestea se încadrează în categoria microîntreprinderilor.



1. **Există o corelație între venituri și cifra de afaceri pe datele înregistrate ? Dacă da, ce tip de corelație ?**

Formula predicției: **Veniturile Totale= 1.0581\* Cifra de Afaceri + 37056.7**

**Venituri** – variabila independentă

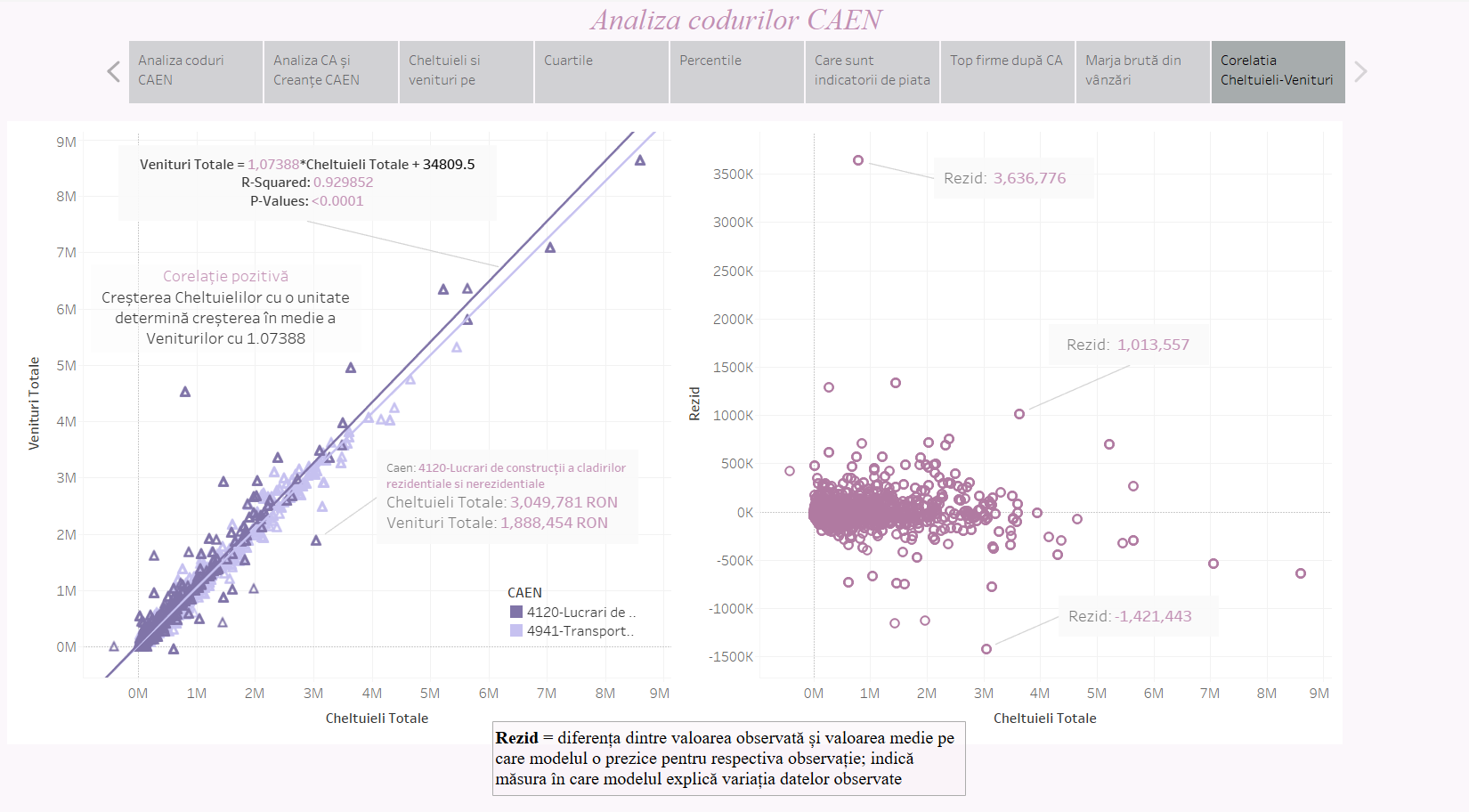
**Cifra de afaceri** – variabila dependentă

O schimbare a variabilei independente cheltuieli totale va rezulta o crestere a veniturilor, deci este vorba de o corelație pozitivă între cele două.

**R-Squared**: proporția în care variația variabilei dependente este explicată de variabila independentă. 87.28% pentru CAEN 4120 și 99.01% pentru CAEN 4941, înseamnă că un procent de aproximativ 88% reprectiv 91% din mișcarea variabilei dependente sunt explicate de variabila independentă venituri.

**P-values:** 0.0001 = corelație semnificativă statistic între parametrii analizați, deci ipoteza nulă se respinge cu o încredere 99.9%. Astfel se deduce faptul că acest model este reprezentativ însă pentru o maibună verificare se construiește residuals plot.

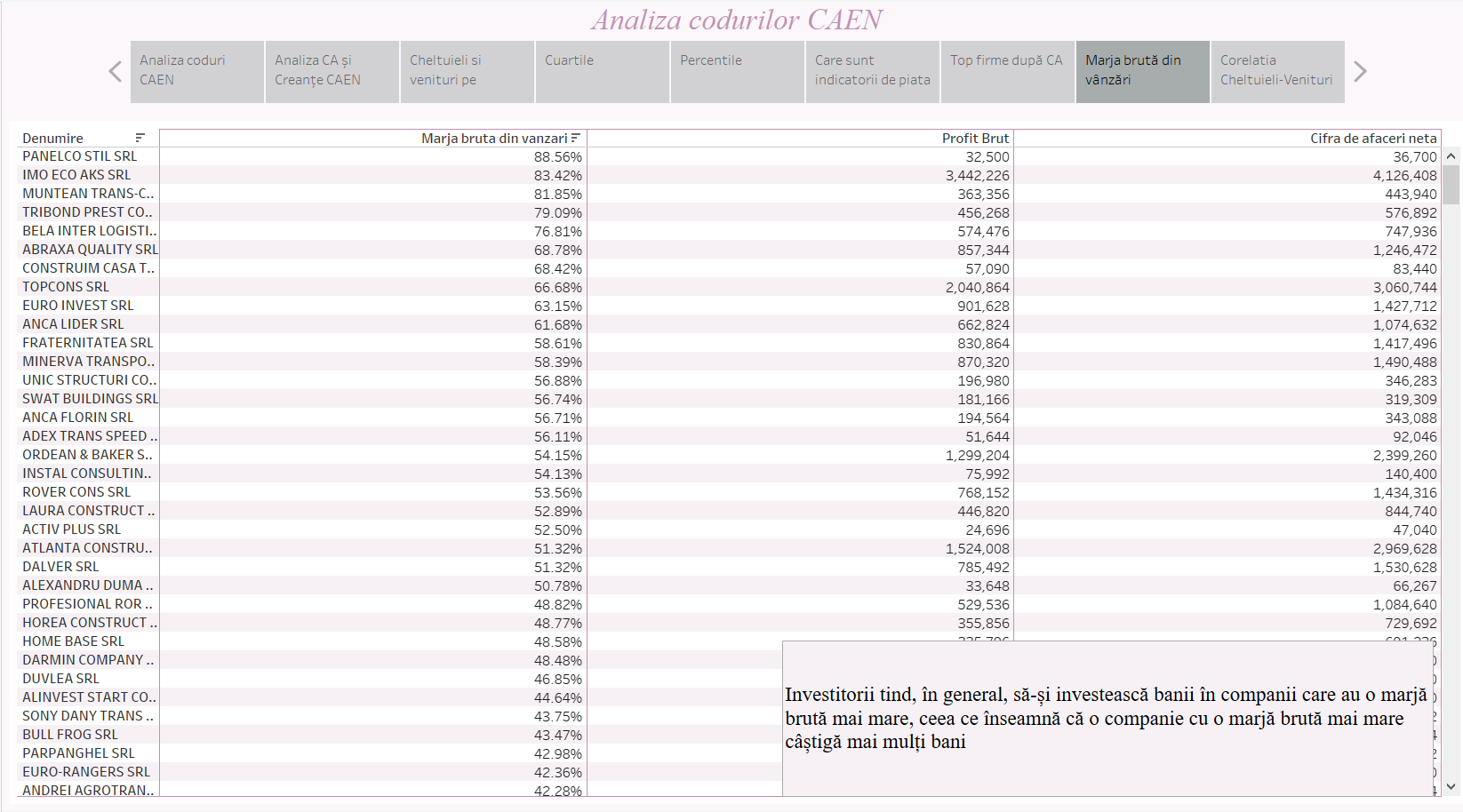
Distribuție normală în jurul valorii zero cu erori reziduale de 3 636 776 = nu contine variabile esentiale, astfel încât se exclud valorile cu o abatere foarte mare.





1. **Care este marja brută din vânzări a firmelor ?**

Pentru a putea răspunde la această întrebare a fost nevoie de calcularea câmpului Marja brută din vânzări cu ajutorul căruia am reușit să analizăm fiecare firmă în parte. Marja brută reprezintă procentajul din totalul vânzărilor reținute de societate după ce toate costurile asociate cu producerea și vânzarea de bunuri și servicii au fost contabilizate. Se dorește ca această marjă să fie cât mai mare, deoarece asta înseamnă un profit brut mai mare.



1. **Concluzie: care este cel mai bun domeniu de activitate dintre cele 2 CAEN-uri pentru a deschide o firmă în anul 2021 ?**

În concluzie, în urma analizei făcute mai sus, codul CAEN cel mai reprezentativ pentru județul analizat este 4941 - Transporturi rutiere de mărfuri deci se recomandă studierea domeniului în amănunt și încercarea implementării unei afaceri pe anul curent în acest sector de activitate.