

AutoBridge Analytics: Data Methodology

მონაცემთა მეთოდოლოგია (Georgian)

1. ფილოსოფია და გამჭვირვალობა

სიზუსტე და აბსოლუტური გამჭვირვალობა AutoBridge Analytics-ის საფუძველია. დარწმუნებული ვართ, რომ საქართველოს საავტომობილო ბაზარს, როგორც რეგიონში ყველაზე დიდ რეექსპორტის ჰაბს, სჭირდება საიმედო ციფრული ინფრასტრუქტურა. ჩვენი სტატისტიკა არ ეფუძნება გამყიდველთა გამოკითხვებს ან ექსპერტების ხელით შეფასებებს, არამედ წარმოადგენს რეალური ციფრული საავტომობილო ბაზრის პირდაპირ, უწყვეტ წყაროს.

2. მონაცემთა შეგროვება (Data Collection)

ჩვენი ETL (Extract, Transform, Load) ალგორითმები ყოველდღიურად აგროვებენ მონაცემებს საქართველოს უმსხვილესი საავტომობილო მარკეტპლეისებიდან და პლატფორმებიდან.

- მოცულობა:** ჩვენი მონაცემთა ბაზა ყოველთვიურად ივსება ათიათასობით ახალი უნიკალური შეთავაზებით.
- სიზუსტე:** მონაცემები კონტროლდება რეალურ დროში, რაც საშუალებას გვაძლევს დავაფიქსიროთ ფასების კლება ან ზრდა, გარბენის ცვლილება და პუბლიკაციის ფინალური სტატუსი (გაყიდული/ნაშლილი).

3. დასუფთავება და ვერიფიკაცია (ETL Pipeline)

ნედლი მონაცემები ყოველთვის სავსეა „ხმაურით“, დუბლიკატებითა და სპამით. ჩვენ ვიყენებთ მკაცრ საინჟინრო ფილტრებს:

- დედუქლიკაცია:** მგა-დილერები ხშირად აქვეყნებენ ერთი და იგივე ავტომობილს სხვადასხვა საიტზე ან ქმნიან დუბლირებულ განცხადებებს თარიღის გასაახლებლად. ჩვენი სკრიპტები აერთიანებენ ასეთ პუბლიკაციებს: ჩვენ ვითვლით უნიკალურ ავტომობილებს და არა მათი სარეკლამო განთავსების რაოდენობას.
- ანომალიები და სპამი:** განცხადებები ფასით 1\$ ან 999,999\$, ან ძველი ავტომობილებისთვის მითითებული 1 კმ გარბენით ავტომატურად გამოირიცხება ანალიტიკური შერჩევიდან.
- სეგმენტების იზოლაცია:** მსუბუქი ავტომობილების მეტრიკაში ჩვენ ალგორითმულად გამოვრიცხავთ სატვირთოებს, სპეცტექნიკას, დანაწილებულ (დაშლილ) ავტომობილებს ნაწილებისთვის და დღიურ ქირაობას.

4. საკვანძო მეტრიკები

- მედიანური ფასი (Median Price):** მოდელის ღირებულების შესაფასებლად ჩვენ არასდროს ვიყენებთ საშუალო არითმეტიკულს (Average). ერთმა შეცდომით ძვირმა ან დაზიანებულმა ავტომობილმა შეიძლება ძლიერად დაამახინჯოს საშუალო ფასი. მედიანა აჩვენებს რეალურ ფასს, რომლითაც იყიდება "ბაზრის შუაგული".

- **ლიკვიდურობის ინდექსი (Liquidity Index):** ეს უნიკალური მაჩვენებელი ასახავს პლატფორმაზე განცხადების სიცოცხლის მედიანურ დროს (დღეებში). რაც უფრო დაბალია ინდექსი (ნაკლები დღე), მით უფრო სწრაფად იყიდება ავტომობილი, რაც ცხელ მოთხოვნაზე მიუთითებს.
- **ბაზრის წილები (Market Share):** განისაზღვრება პროპორციულად, არჩეულ დროის მონაკვეთში კონკრეტული ბრენდის ვალიდური, დუბლიკატებისგან გასუფთავებული პუბლიკაციების მოცულობით.

5. მონაცემთა შეზღუდვები (Limitations)

როგორც ნებისმიერ ანალიტიკურ პლატფორმას, ჩვენც გვაქვს მეტრიკული შეზღუდვები. ჩვენ ვაანალიზებთ ბაზრის "ციფრულ კვალს". გარიგებები, რომლებსაც დილერები ახორციელებენ მთლიანად „ოფლაინში“ (პირდაპირ ავტობაზრობაზე ინტერნეტში გამოქვეყნების გარეშე), არ ხვდება ჩვენს სტატისტიკაში მათ გაციფრულებამდე. თუმცა, ჩვენი ციფრული შერჩევის მასშტაბი უზრუნველყოფს ტრენდების 95%-ზე მეტ სიზუსტეს.