

Dokumentacja projektu

Przetwarzanie dużych zbiorów danych

*Rynek osobowych samochodów używanych oraz nowych
(OtoMoto)*

Autorzy:

Bartosz Kustra 249494

Tomasz Nadrowski 243343

Michał Rudawski 249045

Bartosz Żurkowski 249083

Wrocław, 2021

Spis treści:

1. Wstęp.....	3
2. Opis projektu.....	5
3. Wykorzystywane technologie.....	6
4. Pobrane dane.....	7
5. Analiza danych:.....	8
• Stan: używane oraz nowe.....	8
• Roczniki produkcji.....	8
• Ceny wyjściowe.....	10
• Kraje pochodzenia.....	12
• Rodzaje paliwa.....	14
• Rodzaje skrzyni biegów.....	15
• Rodzaje nadwozia.....	16
• Najpopularniejsze marki.....	18
• Najpopularniejsze modele.....	20
• Najchętniej wybierane kolory lakieru.....	20
• Największa moc silnika.....	22
• Najwyższa cena wyjściowa.....	23
• Najdroższe zabytki.....	25
6. Bibliografia.....	28

1. Wstęp

Spośród oferowanych w obrębie kursu Przetwarzanie Dużych Zbiorów Danych, możliwych do realizacji ścieżek projektowych, omawiana w widocznym projekcie jest ta, oznaczona numerem cztery.

Wyróżniającą ją cechą na tle pozostałych możliwości, była opcja poruszenia wybranego przez grupę projektową faktycznego zagadnienia, które – jak w tym przypadku – powiązane byłoby z życiem codziennym, każdej z osób współtworzących pracę. Dzięki wykonanemu projektowi, możliwe było pozyskanie nie tylko niezwykle istotnej wiedzy praktycznej z zakresu przetwarzania danych, ale również zaspokojenie swojej ciekawości związanej z danym obszarem gospodarczym w Polsce (lub też oczywiście w skali globalnej – uwzględniając pozostałe kraje globu).

Wybierając spośród wielu obszarów życia codziennego, spoglądając na pokrewne między członkami grupy projektowej zainteresowania, oczywistą tematyką, nasuwającą się w pierwszej myśli była motoryzacja. Posiadając doświadczenie w obrębie poszukiwań wymarzonych środków transportu, oraz ogólnorozumiane obeznanie z ową sferą życia, zaskoczeniem nie było wybranie serwisu www.otomoto.pl, będącego prawdopodobnie największym w Polsce, dedykowanym w pełni motoryzacji, portalem aukcyjnym. Zapewnia on niezliczoną możliwość konfiguracji, personalizacji własnych potrzeb oraz oczekiwań względem wyszukiwania, a co za tym idzie udostępnia pokąsną listę różnego rodzaju filtrów. Każdy z nich może stanowić osobne pole do analizy, ponieważ udostępnia przeważnie kilka wariantów odpowiedzi na filtrujące „zapytanie”.

Co oczywiste, z uwagi na pragnienie stworzenia jak najbardziej rozbudowanego projektu, pozwalającego na jak najbardziej dogłębną analizę zagadnienia, wykorzystywane były dane pozyskane z filtrów podlegających pod nazwane na stronie internetowej: filtrowanie zaawansowane. Opcja ta przekierowuje użytkownika na podstronę dedykowaną samochodom osobowym, w obrębie których prowadzona będzie analiza.

Widocznym w tytule jest, iż przeprowadzone badania dotyczyły zarówno samochodów używanych, jak i nowych. Pomimo faktu, iż na OtoMoto zamieszczane są wyłącznie ogłoszenia, mające na celu sprzedaż samochodów (lub części samochodowych, lecz to nie jest przedmiotem badań), ogłoszenia ów, mogą zostać zamieszczone zarówno przez osoby prywatne oraz dealerów należących do polskich sieci danych marek.

Niezwykle istotnym aspektem, wpływającym również w znaczącym stopniu, na dokonany wybór ścieżki projektowej oraz poruszonej weń tematyki, była potężna i mocno rozwinięta społeczność serwisu. Wszakże działalność OtoMoto nie ogranicza się jedynie do odpowiedniego zarządzania i utrzymywania strony internetowej. Istotnym jest również kontakt z ludźmi bezpośrednio lub pośrednio zaangażowanymi w ten niebanalny projekt. Pomijając liczne publikacje wideo, dostępne na serwisach YouTube oraz Facebook, które nie są przedmiotem analizy projektowej; istotnym jest udostępnianie szerokokorozumianej publice odpowiednich statystyk, dotyczących serwisu, aukcji nań wystawianych oraz wielu innych aspektów funkcjonowania rynku samochodów używanych oraz nowych w Polsce. Wszystko to zawierane jest w formie raportów, których analizy dotyczą okresu jednego roku kalendarzowego.



Rys 1. Okładka raportu rocznego OtoMoto 2018 – Internetowy Samochód Roku

Kierując się ciekawością, odnoszącą się do wspomnianych już raportów, grupa projektowa postanowiła stworzyć niejako odpowiedź, w postaci własnoręcznie sporządzonego raportu. Co oczywiste, nieunikniona jest znacząca różnica w jakości wspomnianych publikacji, względem pracy studenckiej, co determinowane jest wieloma czynnikami. Pierwszym z nich jest brak możliwości pozyskania informacji z przestrzeni całego roku kalendarzowego – wpływ mają na to ograniczenia sprzętowe oraz, co być może bardziej istotne, możliwość wykrycia przez serwis OtoMoto niepożądanego ruchu sieciowego, który miałby charakter cykliczny i niezwykle częsty (scraping danych). Kolejnym z aspektów uniemożliwiających dokładne odtworzenie „konkurencji” prezentowanych w oficjalnym raporcie, jest fakt, iż przeciętny użytkownik, jakim jest każdy z realizujących zadanie studentów, nie posiada dostępu do informacji czysto administracyjnych, takich jak choćby: ilość użytkowników zarejestrowanych na stronie czy też zlecane wyszukiwania, spełniające postawione przez użytkowników kryteria.

Analiza zatem ogranicza się do powszechnie dostępnych informacji, pozyskanych bezpośrednio ze strony, na której operować i wprowadzać zmiany może przeciętny użytkownik portalu, nie administrator.

2. Opis projektu

Wspomnianą inspiracją do stworzenia projektu były udostępniane przez OtoMoto, cyklicznie – rocznie, raporty zdradzające wiele interesujących informacji odnoszących się zarówno do tendencji widocznych wśród samych konsumentów – potencjalnych lub faktycznych kupujących samochody – oraz także w świecie motoryzacji, dostępne na portalu są bowiem, wspomniane już, ogłoszenia samochodów nowych.

Dostęp do opisywanych publikacji jest całkowicie bezpłatny oraz powszechny. Pobieranie jest możliwe zarówno bezpośrednio ze strony tytułowego portalu internetowego oraz ze źródeł współpracujących i współdziałających w obrębie realizacji raportów: Samar, Carsmile – uzależnione jest to od roku publikacji dzieła.

Głównym celem widocznego projektu, jest zatem skonfrontowanie pozyskanych informacji oraz próba wykonania zbliżonego ideą raportu, którego główna tematyka dotyczyłaby właśnie statystyk z obszaru samochodów osobowych. Jak już zostało wspomniane, pominięte zostały obszary sprzedaży części samochodowych, motoryzacji jednośladowej oraz pojazdów „ciężkich”.

Prawdopodobnie największym wyzwaniem, powiązaniem z wykonywaniem tegoż projektu, było odpowiednie pobranie oraz obrobienie surowych danych, w taki sposób, by później w przejrzysty sposób móc wykorzystywać zasoby programistyczne, w celu dokonywania jak najbardziej wnikliwej analizy. Obszar motoryzacji jest niezwykle przepastny, czego najlepszym przykładem jest skala poddanych analizie 212 149 rekordów, w którym na każdy z nich składały się wartości 32 kolumn.

Po pobraniu wspomnianej – potężnej ilości danych, musiałyby one zostać odpowiednio usystematyzowane w tabelę, wedle przytoczonych wcześniej kolumn, tak by każda dana odpowiadała dla siebie bliźniaczym, w pionowej analizie. Szczególną uwagę należało poświęcić występującym – nielicznym, lecz jednak duplikatom. Obecność ogłoszeń zawierających te same parametry (oczywiście w pełnym zakresie) wymagała pozostawienia zaledwie jednej „kopii” takiegoż anonsu. W przypadku niewystosowania wspomnianego procesu, możliwe byłoby zachwianie i zaburzenie precyzji przeprowadzonej analizy, gdyż niemożliwe byłoby pozostanie rzetelnym wobec odbiorcy.

Zważywszy na ciągle rozwijającą się bazę ogłoszeń na portalu OtoMoto – zarówno pod względem ich liczby oraz dostępnych do ustawiania parametrów (przykładem może się stać dodanie w obrębie możliwości kreacji nowych marek, a tym samym modeli), należy przyjąć, iż prezentowane statystyki mają charakter poglądowy, wszelkie pobrane dane pochodzą z dnia 14 kwietnia 2021 roku oraz ówczesnie były najbardziej aktualną z oferowanych wersji, wszakże OtoMoto udostępnia możliwość edytowania dodanych wcześniej ogłoszeń.

3. Wykorzystywane technologie

Jak zostało to już wspomniane we wcześniejszej części projektu: potężna baza danych uzupełniana o pola odpowiadające konkretnym filtrom OtoMoto, sprawiła, iż przetwarzanie tychże pól byłoby niemożliwe do wykonania w akceptowalnym czasie pracy, połączonych mocy obliczeniowych komputerów wszystkich osób biorących udział w projekcie. Niezbędne było zatem wykorzystanie rozwiązania chmurowego. Dzięki temu, wszystkie zlecone procesy mogły zostać wykonywane, nie zabierając – wspomnianej już – mocy obliczeniowej urządzeń studenckich – niejako pracując w tle. Dodatkowo, przetwarzanie danych, z wykorzystaniem operacji przeprowadzanych w chmurze, następowały szybciej, niż miałyby to miejsce w przypadku konwencjonalnego podejścia.

Wybór systemu chmurowego był dokonywany pomiędzy dwoma, prawdopodobnie najbardziej popularnymi obecnie rozwiązaniami: AWS oferowane przez Amazon lub Google Cloud. Każdy ze studentów biorących udział we współtworzeniu projektu, na innych zajęciach – w tym także laboratoryjnych – miał już styczność z AWS, zatem wybór padł na rozwiązanie oferowane przez Google. Celem tegoż działania było stworzenie możliwości i sposobności do zapoznania się z nowym środowiskiem oraz poszerzenia swoich horyzontów w obszarze rozwiązań rozproszonych.

Następnie wyekstrahowane już dane, należało w odpowiedni sposób oczyścić, tak by wyeliminować wszelkie niedopatrzenia, poczynione przez osoby wystawiające ogłoszenia, w trakcie ich publikowania. Najczęstsze błędy dotyczyły formatu zapisu ceny, który był niejednoznaczny oraz niespójny z innymi udostępnianymi. Zagadnienie to zostało ułatwione dzięki oferowanym przez Pandas. Jest to biblioteka stworzona do współpracy z aplikacjami tworzonymi w języku programowania Python. Nazwa wspomnianej biblioteki, została zaproponowana w oparciu o połączenie słów we frazie: „panel data”, które jest pojęciem odnoszącym się bezpośrednio do analizy danych. Dodatkowo, niektóre źródła, podają, iż nazwa powiązana jest również z: „Python data analysis”. Niezależnie jednak od etymologii nazwy samej biblioteki, jest ona niezwykle przydatna w projektach, przypominających lub będących w większej części spokrewnionych z prezentowanym. Ogromna liczba funkcjonalności pozwala między innymi na: indeksowanie danych, dokonywanie zmian w obrębie poszczególnych kolumn (ich dodawanie lub też usuwanie), łączenie poszczególnych porcji danych czy też najważniejsze z perspektywy tegoż projektu: odpowiednie filtrowanie danych. Po wspomnianym „oczyszczeniu” pobranych danych, należało wykorzystać ostatnią z wymienionych właściwości Pandas, by móc zobrazować w przyszłości pożądane wyniki przeprowadzanej analizy.

Oprócz biblioteki Pandas, w części programowej projektu, zaimplementowane zostały również biblioteki: Numpy – pozwalająca na analizę wielowymiarowych wolumenów danych oraz Matplotlib, dzięki której możliwe było kreślenie odpowiednich wykresów, które w „namacalny” sposób przedstawiałyby analizy przeprowadzone w obrębie pobranych danych.

4. Pobrane dane

We wcześniejszej części projektu, wspomniana została potężna ilość niezbędnych do przeanalizowania danych. Należałoby zatem pochylić się nad tym, w jaki sposób, informacje, które widoczne są w ogłoszeniach zamieszczonych w tytułowym portalu motoryzacyjnym, znalazły swoje miejsce w zapisie, umożliwiającym ich dalszą analizę.

Całość części programistycznej, opierała się zatem na dwóch skryptach, które były ze sobą w odpowiedni sposób połączone, tak by zapewnić w ogólnym rozrachunku pełne pokrycie danych, pobieranych z serwisu internetowego. Pierwszy z nich, determinował podwalinę pod resztę pobieranych danych, klasyfikując i porządkując marki samochodów, które w momencie scrapingu były oferowane, jako możliwości wyboru w trakcie kreacji ogłoszeń. Obecnie, stale analizując tytułowy serwis internetowy, widocznym jest, iż poczynione zostały kolejne ekspansje, a zatem przybyło kolejnych marek o raz – idąc w sukurs – również i modeli.

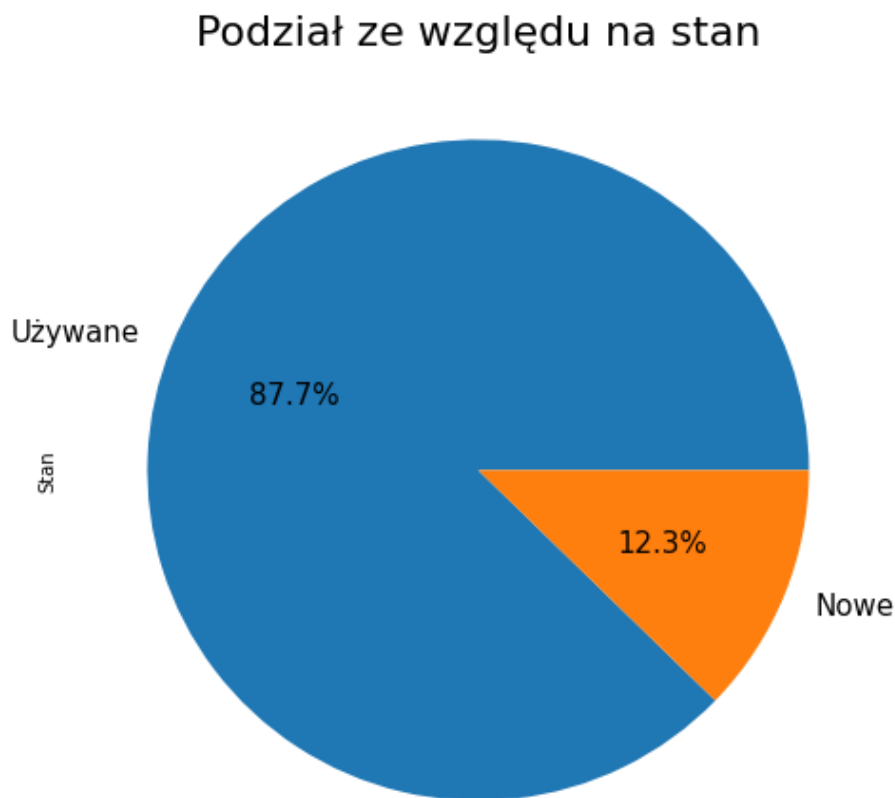
Drugi natomiast skrypt, pozwalał na przyporządkowanie do konkretnych marek, podlegających poden modeli oraz opisanego w ogłoszeniu:

- Parametry powiązane z lokalizacją oferowanego pojazdu,
- Parametry powiązane z silnikiem,
- Cena,
- Kraj pochodzenia,
- Status pojazdu:
 - Czy jest zarejestrowany w Polsce,
 - Czy jest bezwypadkowy,
 - Czy jest zarejestrowany jako zabytek,
 - Czy posiada kierownicę po prawej stronie (Anglik),
 - Czy był serwisowany w Autoryzowanym Serwisie,
- Czy posiada filtr DPF,
- Rok produkcji,
- Data pierwszej rejestracji pojazdu,
- Przebieg,
- Rodzaj oferty i finansowania,
- Parametry powiązane z napędem,
- Typ nadwozia,
- Kolor lakieru,
- Stan pojazdu (nowy lub używany),
- Czy dany pojazd jest obecnie w rękach swojego pierwszego właściciela,
- Data dodania ogłoszenia w serwisie,
- Link do ogłoszenia.

Tak pobrane i powiązane ze sobą dane, zostały umieszczone w tabelach: z podziałem ze względu na markę oraz w zbiorczej. Te pierwsze pomagały w weryfikacji poprawności funkcjonowania oprogramowania, druga natomiast stała się bazą pozwalającą na dokonywanie wszystkich analiz zamieszczonych w poniższym rozdziale.

5. Analiza danych

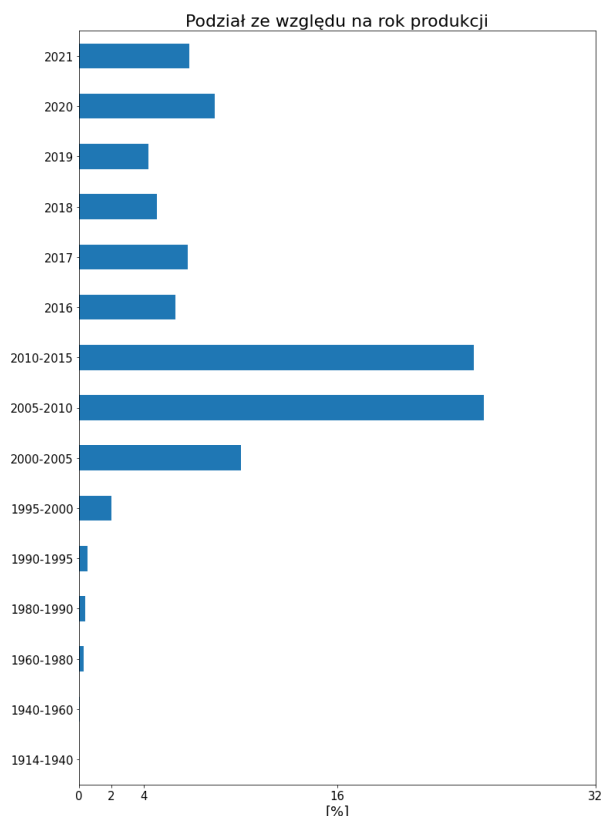
Pierwszym z zagadnień, które należałoby poruszyć jest oczywiście kwestia stanu samochodów, których dotyczą ogłoszenia na OtoMoto – jak rozkłada się liczba aut nowych oraz używanych. Wykres zaprezentowany został poniżej:



Rys 2. Procentowy podział samochodów ze względu na ich stan: nowy lub używany (kwiecień 2021)

Co jest więc widoczne: pomimo panujących trendów, w których to coraz więcej autoryzowanych serwisów zaczyna wdrażać się w rynek internetowy, za pośrednictwem portalu OtoMoto, wciąż znacznie mocniej dominują na rynku samochody używane. Wiele informacji z branży motoryzacyjnej wskazuje, iż stosunek tych dwóch wielkości będzie dążył do redukcji panującej między nimi różnicy, są to jednak prognozy rozplanowane na wiele najbliższych lat, nie należy zatem traktować ich, jako pewnych.

Wiedząc już, że zdecydowanie bardziej popularne w serwisie są ogłoszenia, których przedmiotem są samochody używane, interesującym zagadnieniem byłoby, dowiedzieć się, jak prezentuje się rozkład roczników produkcji danych maszyn:



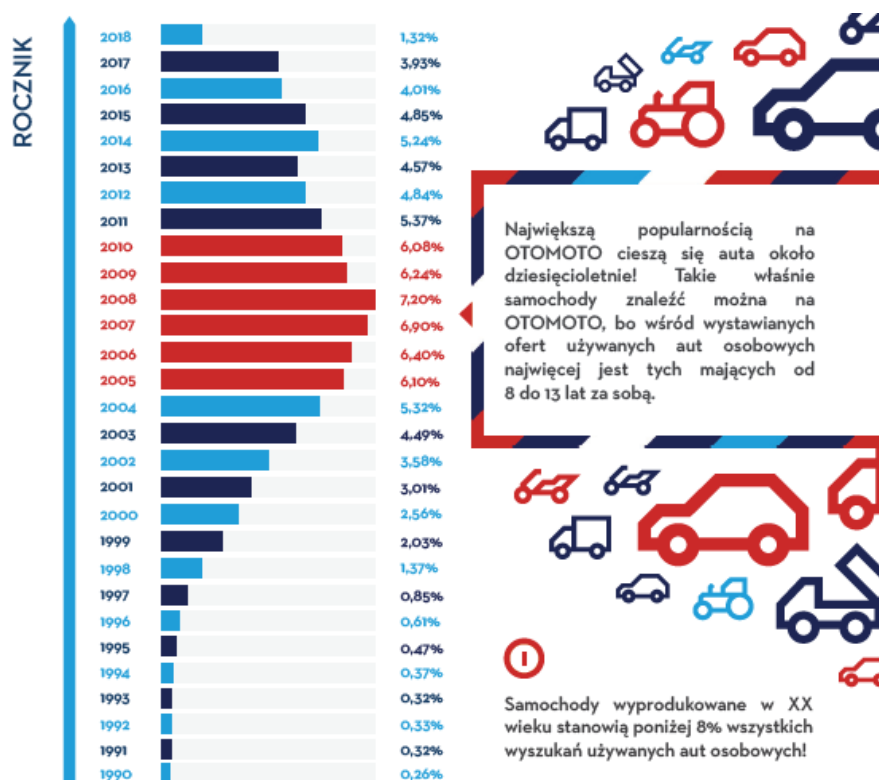
Rys 3. Rozkład samochodów ze względu na rok produkcji (kwiecień 2021)

Oprócz najbardziej oczywistych wniosków, jakoby samochody, które zostały wyprodukowane w latach 2005- 2015 zdominowały rynek (wszakże jest to też największe pole czasowe), interesującym jest, że dostępnych jest więcej samochodów z obecnego roku (pomimo faktu, że raport sporządzany był w kwietniu), niż z lat: 2019 czy 2018. Najprawdopodobniej zatem, gdy rok kalendarzowy będzie się chylił ku końcowi, przebita może zostać wartość modeli wyprodukowanych w roku 2020.

Opisany powyżej proces, może stanowić niejako odpowiedź na rosnące zapotrzebowanie na samochody młodsze, czego dowodzić mogą statystyki sporządzone przez OtoMoto i udostępnione w najnowszym raporcie corocznym (2020), które dotyczą poszukiwań użytkowników serwisu:



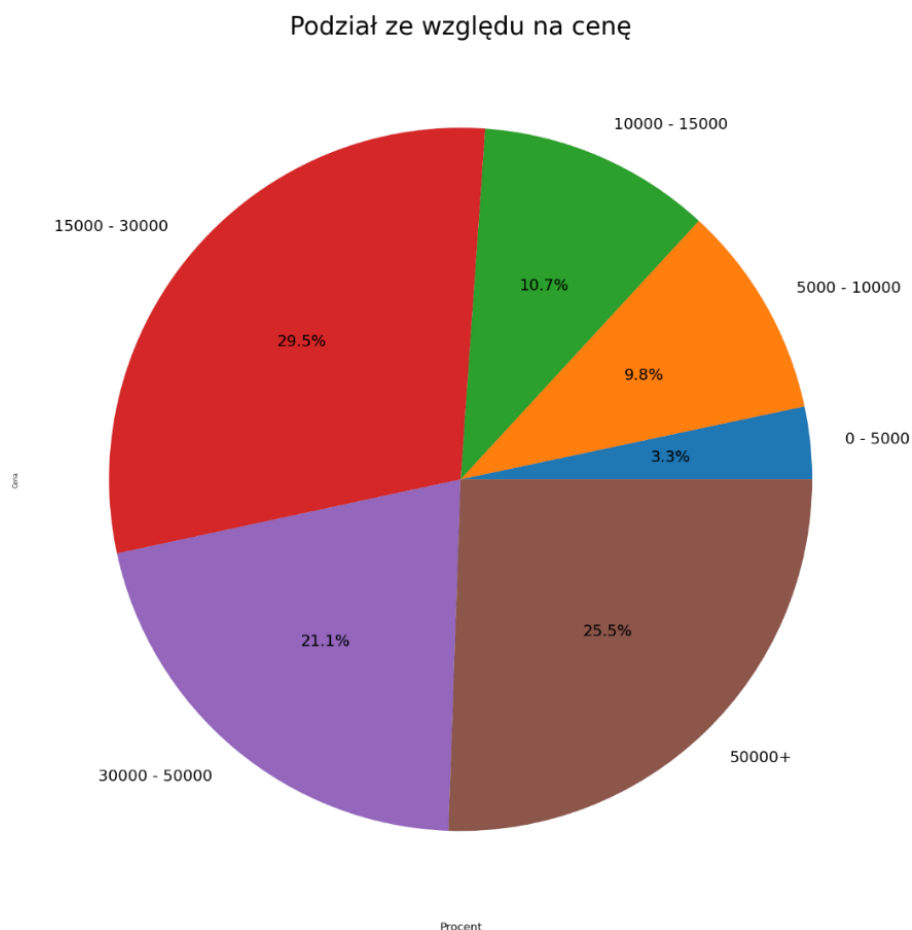
Rys 4. Procent wyszukiwań danego rocznika (OtoMoto 2020)



Rys 5. Procent wyszukiwań danego rocznika (OtoMoto 2018)

Można zatem stwierdzić, że poszukiwane samochody i stawiane względem nich oczekiwania, znajdują pewne odzwierciedlenie w zamieszczanych ogłoszeniach. Ponadto na powyższe wykresy składać może się między innymi czynnik, mówiący o przeprowadzanej „zmianie pokoleniowej” wśród użytkowników dróg. Poszukiwane są nowsze samochody, jednocześnie – wystawiane zostają dotychczasowe automobile. Przyczynia się to do zamknięcia obiegu rynku samochodów używanych, a jednocześnie implikuje wzmożony ruch w świecie nowej motoryzacji. Na tle Europy Polska jawi się jako jeden z krajów, w których średni wiek samochodu jest najstarszy. Prezentuje się on na poziomie 13,6 lat, a dla porównania w Niemczech na poziomie 9,3 roku. Tendencja jednak, wskazuje, iż Polacy dążą do posiadania jak najnowszych samochodów, ale również i do częstszego ich wymieniania.

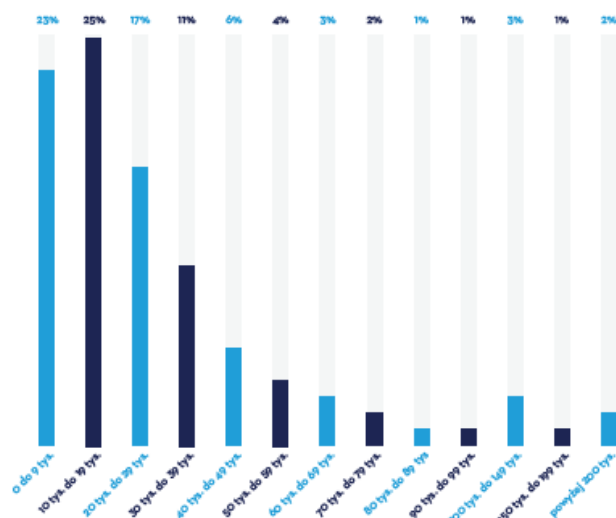
Posiadając wiedzę dotyczącą roczników wystawionych w ogłoszeniach samochodów oraz tych najczęściej wyszukiwanych w poprzednich latach, zastanawiające mogą być najbardziej popularne przedziały cenowe, w których oferowane (oraz poszukiwane) są pojazdy:



Rys 6. Procentowy rozkład samochodów ze względu na ich cenę (kwiecień 2021)

Niezmiennie na przestrzeni lat, najczęstszym przedziałem cenowym, oferowanych pojazdów na platformie OtoMoto, jest od piętnastu do trzydziestu tysięcy złotych polskich. Bardziej intrygującym jest, być może, druga pozycja w tym zestawieniu, a więc samochody osobowe droższe niż 50 tysięcy. Pokrywać się to może ze wcześniej wysnutymi wnioskami, mówiącymi o tym, iż Polacy coraz chętniej sięgają po nowsze samochody, co implikuje wyższą cenę zakupu.

Poniżej zamieszczone zostały wykresy, zaczerpnięte ze wspomnianych raportów rocznych (2018 oraz 2020) OtoMoto: Internetowy Samochód Roku, a które dotyczą statystyk poszukiwań samochodów osobowych, z uwzględnionym podziałem ze względu na ich cenę wyjściową, zamieszczoną w ogłoszeniu.



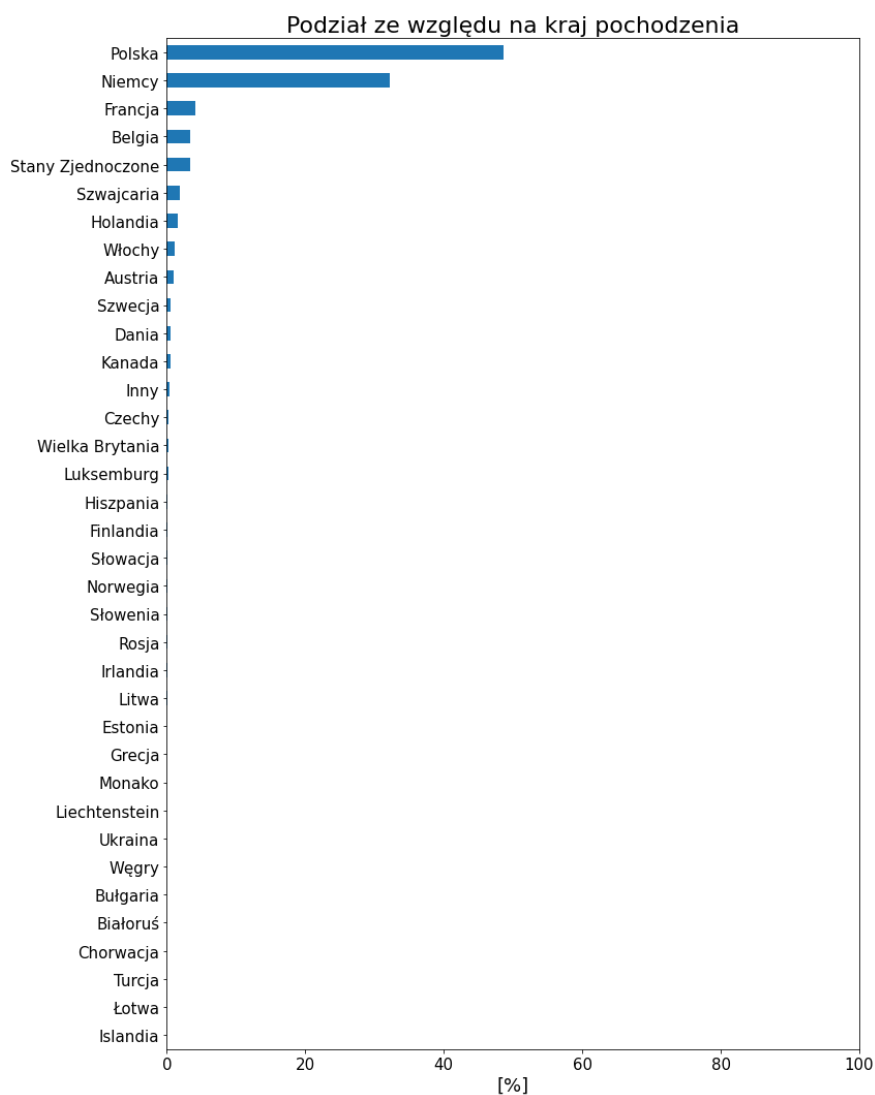
Rys 7. Procentowy podział wyszukiwań samochodów ze względu na ich cenę (OtoMoto 2018)

SAMOCCHODY W JAKICH PRZEDZIAŁACH CENOWYCH SPRZEDAWAŁY SIĘ NAJLEPIEJ?

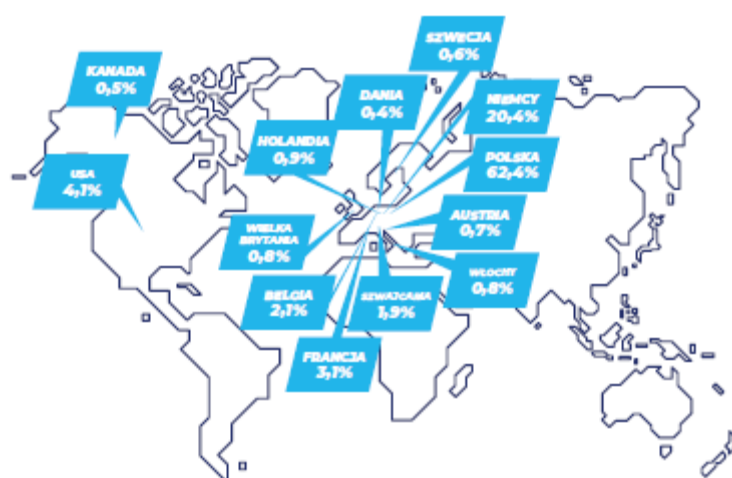


Rys 7. Procentowy podział wyszukiwań samochodów ze względu na ich cenę (OtoMoto 2020)

Zanim jednak jakiegokolwiek auto zostanie wystawione na portalu ogłoszeniowym, musi zostać sprowadzone z innego kraju, lub wydystribuowane z polskiego salonu. Poniżej przeanalizowane zostały kraje pochodzenia wystawionych na sprzedaż samochodów zarówno w ujęciu stworzonym na potrzeby projektu (kwiecień 2021) oraz raportu za rok 2020:

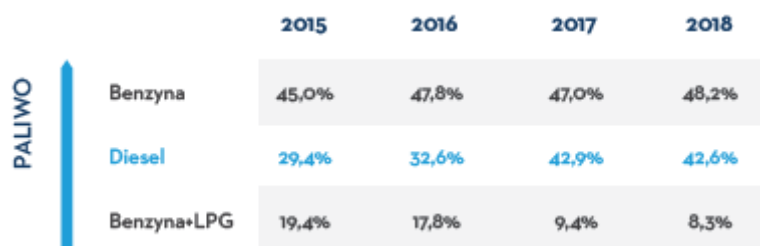


Rys 8. Rozkład procentowy samochodów ze względu na kraj ich pochodzenia (kwiecień 2021)



Rys 9. Procentowy podział wyszukiwań samochodów ze względu na kraj ich pochodzenia (OtoMoto 2020)

Pochylić należałoby się również nad upodobaniami Polaków, odnoszących się do preferowanego rodzaju paliwa napędowego. Poniżej zaprezentowana została statystyka umieszczona w raporcie za rok 2018:



	2015	2016	2017	2018
Benzyna	45,0%	47,8%	47,0%	48,2%
Diesel	29,4%	32,6%	42,9%	42,6%
Benzyna+LPG	19,4%	17,8%	9,4%	8,3%

Rys 10. Procentowy podział wyszukiń samochodów ze względu na wykorzystywane paliwo (OtoMoto 2018)

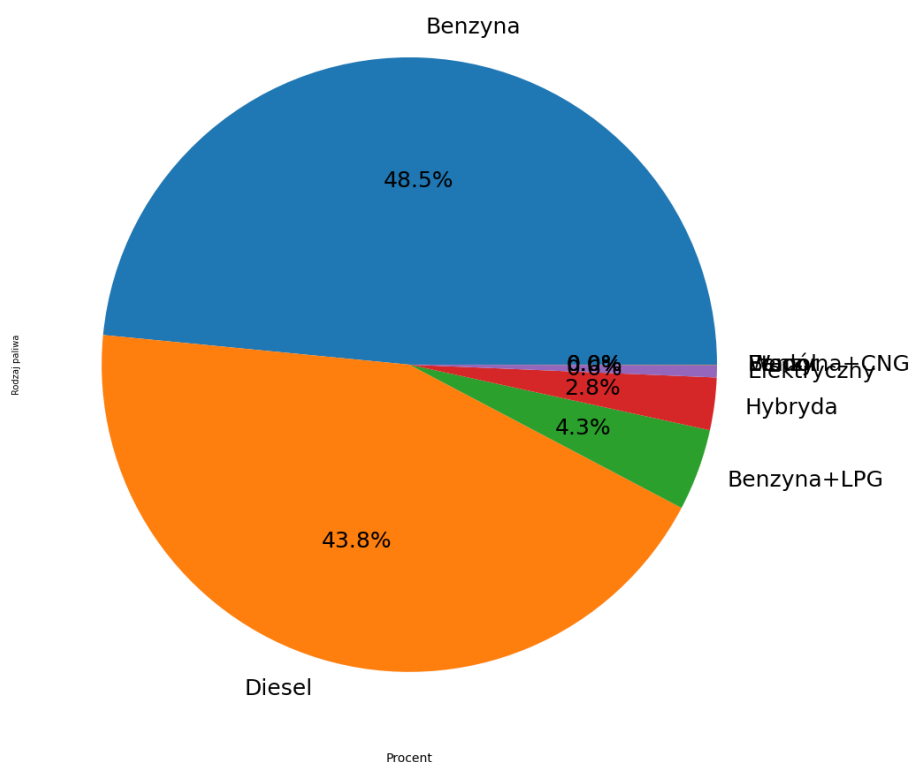
Z powyższej infografiki można wywnioskować, iż z biegiem lat wśród polskiej klienteli zwiększa się zainteresowanie silnikami wysokoprężnymi. Może to wynikać z coraz bardziej dopracowanych konstrukcji silników Diesla, których podzespoły cechują się większą trwałością, a potencjalne naprawy nie muszą oznaczać znacznego ubytku zasobów finansowych. Powiązane może to być również z tym, iż z biegiem czasu Polacy pokonują średnio więcej kilometrów w ciągu jednego roku, a by uczynić to bardziej opłacalnym, logicznym wyborem w trakcie poszukiwania środka transportu, jest skierowanie się w stronę wspomnianego rodzaju silnika, ponieważ cechuje się on przeważnie wyższą sprawnością, niżli ma to miejsce w przypadku napędu benzynowego.

Dodatkowym czynnikiem wpływającym na wzrastającą liczbę poszukiwanych samochodów z silnikiem diesla, jest zwiększony import takowych modeli zza granicy. Powiązane jest to z obecnie panującą sytuacją polityczną i wprowadzanymi – w wielu państwach, dokładnie w centrach większych miastach – zakazami poruszania się pojazdami z silnikami wysokoprężnymi. Naturalnym rynkiem zbytu dla wszystkich tych pojazdów jest Polska.

Silniki benzynowe cieszą się niemalże stałą popularnością, niezależnie od biegu lat, zmiany natomiast dotyczą zainteresowania najbardziej popularnym paliwem alternatywnym, a więc gazem: LPG. Ciężkim w ocenie jest stwierdzenie jednoznacznie, jaka jest przyczyna takiej sytuacji.

Poniżej zamieszczona została statystyka umieszczonych na portalu OtoMoto w perspektywie wykresu kołowego, z uwzględnieniem wszystkich dostępnych w filtrach rodzajów paliwa:

Podział ze względu na rodzaj paliwa



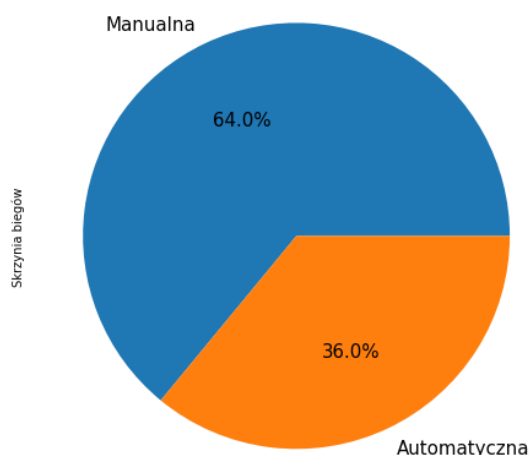
Rys 11. Rozkład procentowy samochodów ze względu na wykorzystywane paliwo (kwiecień 2021)

Widocznym jest zatem, iż w przeciągu kilku lat od udostępnionej w raporcie statystyki, nie zmieniło się wiele. Co oczywiste wcześniej zamieszczane infografiki odnosiły się do poszukiwanych na portalu samochodów, powyżej natomiast przedstawione są ogłoszenia.

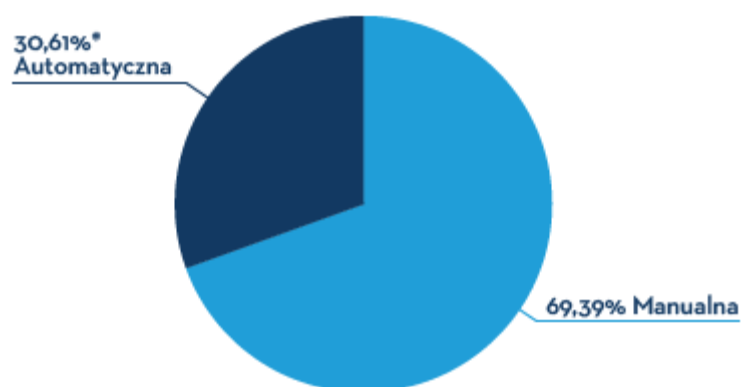
Niemniej jednak, możliwe jest po raz kolejny utwierdzenie się we wniosku, iż zachodzi mocna więź, relacja pomiędzy popytem oraz podażą – rynek dostosowuje się do zapotrzebowania.

Możliwe byłoby zatem pochylenie się nad kolejnym aspektem rynku samochodów osobowych, czyli przekładni zmiany biegów. Od wielu lat upodobania polskich kierowców są kierowane zdecydowanie bardziej w stronę przekładni manualnych, które – zważywszy na większe skomplikowanie konstrukcyjne – mogą odstraszać potencjalnych nabywców, ponieważ związane są z nimi wyższe koszty eksploatacji.

Podział ze względu na skrzynie biegów



Rys 12. Rozkład procentowy samochodów ze względu na skrzynię biegów (kwiecień 2021)

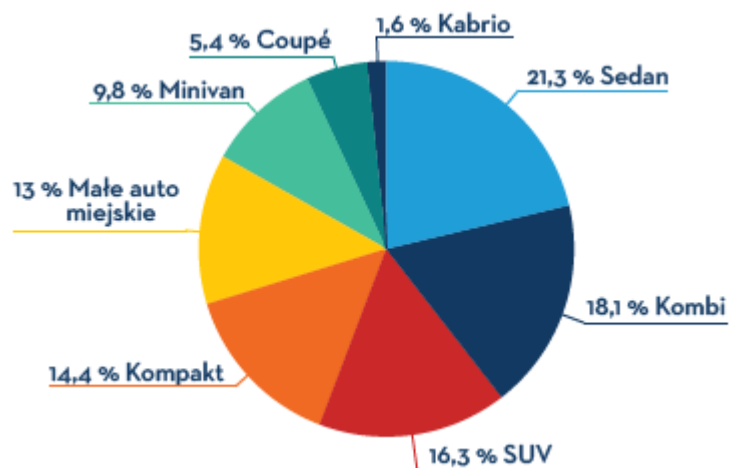


*W tym skrzynie półautomatyczne, których w wyszukiwaniach jest poniżej 1%

Rys 13. Rozkład procentowy wyszukiwań samochodów ze względu na skrzynię biegów (OtoMoto 2018)

Ponownie zestawione zostały: ogłoszenia znajdujące się w kwietniu 2021 roku oraz oczekiwania – poszukiwania użytkowników z roku 2018. Niemniej jednak, wartości procentowe są niezwykle zbliżone.

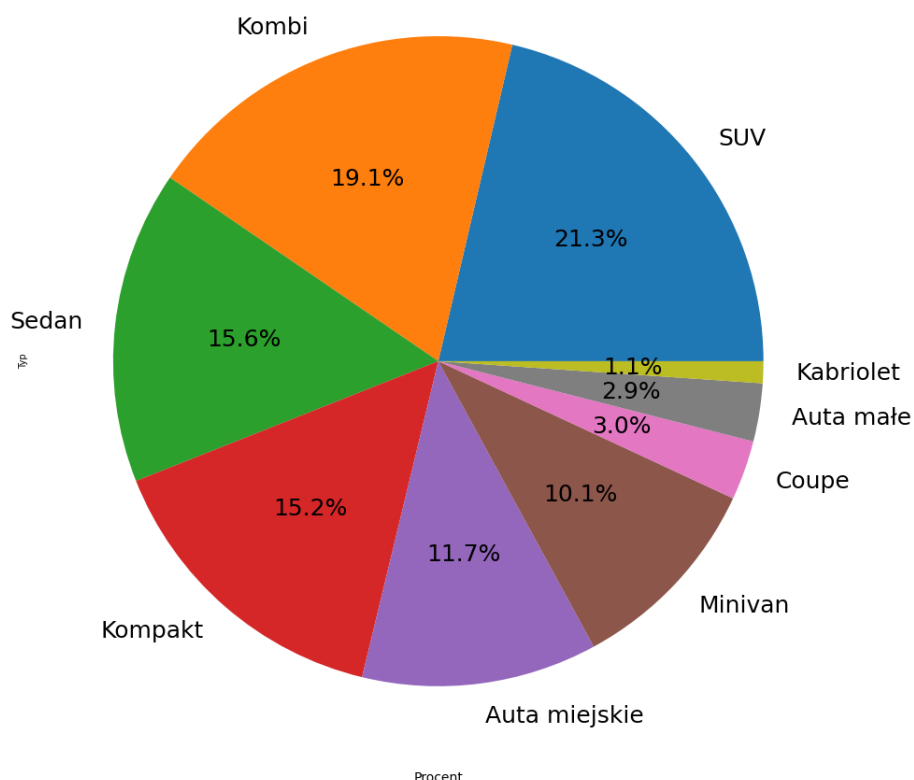
Jakie natomiast typy nadwozia cieszą się największą popularnością wśród polskiej klienteli? Odpowiedź na to odnaleźć można w raporcie udostępnionym w roku 2018. Dwa lata później, statystyki ów dotyczyły już wyłącznie zestawienia dwóch typów karoserii: sedanów oraz SUVów.



Rys 14. Rozkład procentowy wyszukiwań samochodów ze względu na rodzaj nadwozia (OtoMoto 2018)

Tak jak miało to miejsce w przypadku poprzednich, rozpatrywanych kategorii tak i tutaj na powyższej grafice zamieszczone zostały informacje wystosowanych zapytań w portalu OtoMoto. Poniżej natomiast statystyka oferowanych, dostępnych egzemplarzy w kwietniu roku 2021:

Podział ze względu na typ nadwozia



Rys 15. Rozkład procentowy samochodów ze względu na rodzaj nadwozia (kwiecień 2021)

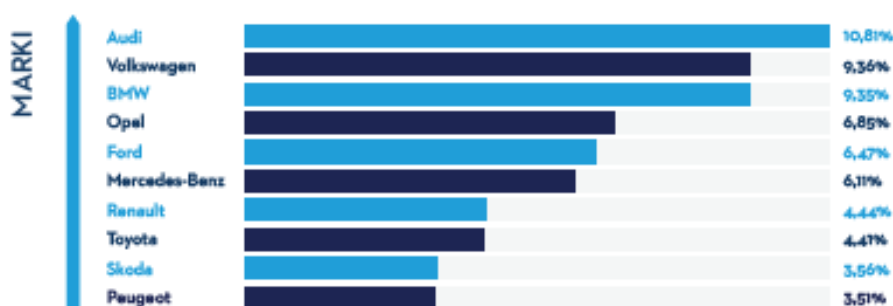
Widocznym jest zatem, który typ nadwozia zdominował ofertę strony internetowej, na przestrzeni ostatnich lat. Zarówno w kwestii poszukiwań oraz oferty rynkowej: SUV staje się swego rodzaju wyznacznikiem obecnych trendów.

Warto jednak podkreślić, iż w przypadku wspomnianego rodzaju nadwozia, ukryta jest pewna luka filtrowania – w obrębie wspomnianej kategorii mogą zostać umieszczone pojazdy z różnych segmentów: począwszy od B, aż po niezwykle luksusowe ponad pięciometrowe kolumbryny, czego nie można powiedzieć np. o samochodach, opisywanych jako „kompakty”. Kategoria ta bowiem – już z samej definicji – determinuje pewne ramy wielkości nadwozia.

Nie zmienia to jednak faktu, iż ekspansja SUVów ma miejsce nie tylko w Polsce, Europie ale również i w Ameryce oraz Azji. Bryła ta najwidoczniej idealnie dostosowała się do oczekiwań ludzi z całego świata. Nie powinno to dziwić, ponieważ sama etymologia nazwy tegoż rodzaju pojazdów: Sport Utility Vehicle – z założenia powinien być to więc środek transportu, który powinien satysfakcjonować niemalże całe spektrum potencjalnych nabywców.

Uwydatniony jest również trend mówiący o tym, iż klienci coraz chętniej decydują się na zakup samochodów większych. Obrazem tego jest nagłe załamanie popularności samochodów małych, miejskich. Można zatem stwierdzić, że z biegiem lat coraz to ważniejszym aspektem – uwzględnianym przy zakupie samochodu – jest zapewnienie sobie większego komfortu oraz możliwości transportu większej objętości bagażów. Wzrost popularności SUVów implikuje również spadek zainteresowania oraz dostępności na rynku wtórnym sedanów, które przez wiele lat stanowiły główną siłę napędową opisywanego obszaru biznesowego. Wynikało to – co oczywiste – z wieloletniej tradycji produkowania samochodów w takim właśnie typie nadwozia.

Skoro wiadomo już, jakie typy karoserii są najpopularniejsze – zarówno wśród poszukiwań, jak i wśród oferowanych ogłoszeń, możliwe jest pochylenie się nad konkretnymi markami. Tymi, które dla Polaków są najbardziej atrakcyjne (ilość wystosowanych zapytań, dla konkretnych filtrów):

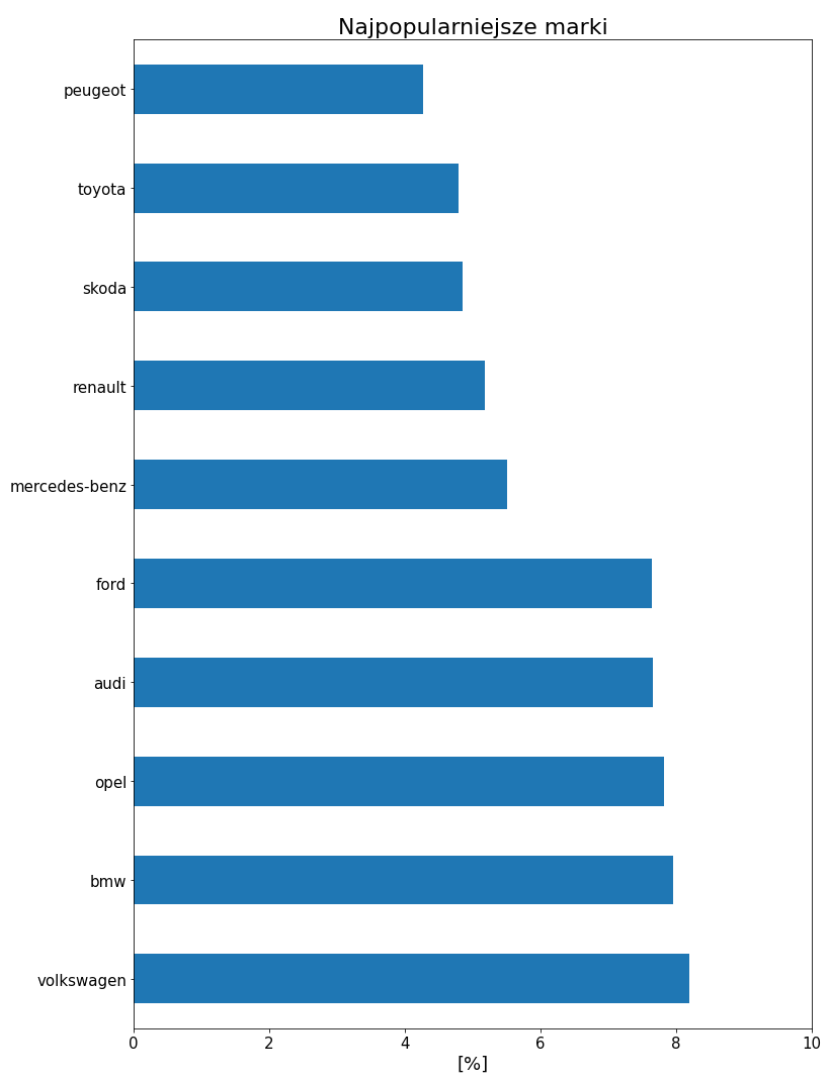


Rys 16. Najczęściej wyszukiwane marki samochodów (OtoMoto 2018)

MARKA	UDZIAŁ
AUDI	11,05%
BMW	10,30%
VOLKSWAGEN	8,43%
MERCEDES-BENZ	7,72%
FORD	6,72%
OPEL	5,63%
TOYOTA	4,24%
RENAULT	3,97%
ŠKODA	3,54%
PEUGEOT	3,24%

Rys 17. Najczęściej wyszukiwane marki samochodów (OtoMoto 2020)

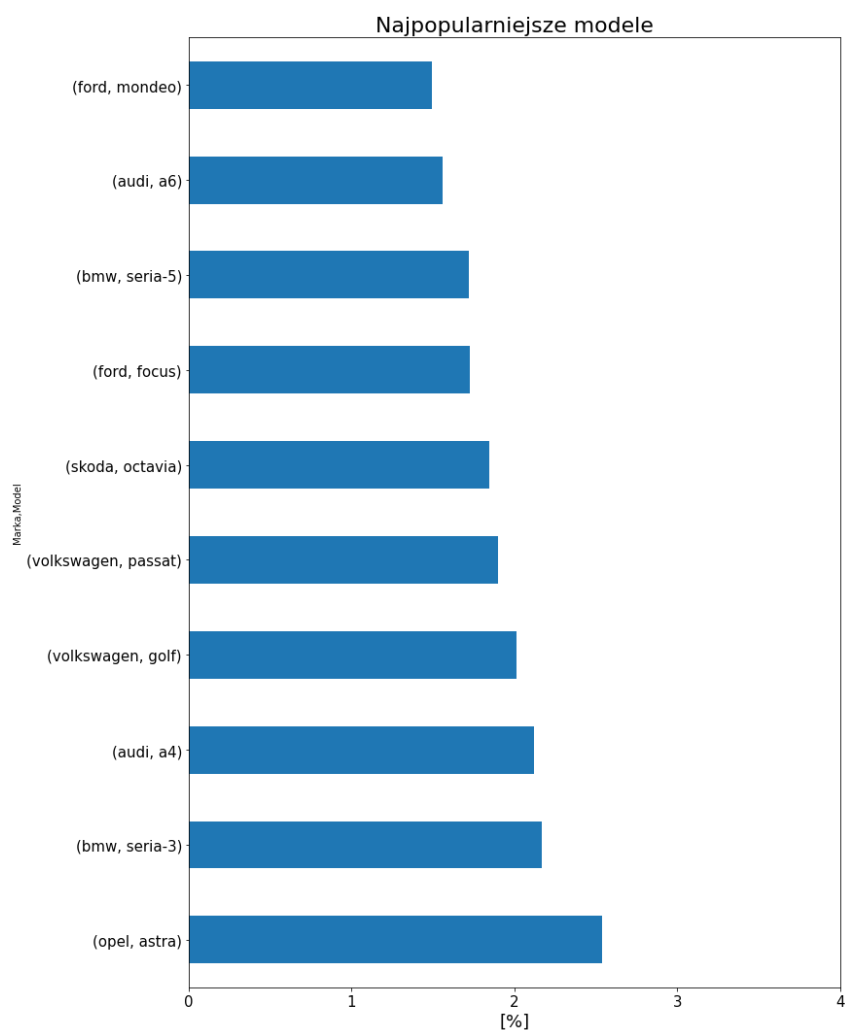
Oraz tych marek, które najczęściej oferowane są przez użytkowników na portalu OtoMoto:



Rys 18. Najczęściej oferowane marki samochodów (kwiecień 2021)

Wyciągając wnioski na podstawie połączenia informacji zawartych w powyższych wykresach – można stwierdzić, że występuje spory dysonans pomiędzy oczekiwaniami względem rynku samochodów używanych, a tym co może on faktycznie zaoferować. Polacy poszukują często niemieckich samochodów marek premium, podczas, gdy rynek przesycony jest tymi mniej prestiżowymi

Najczęściej oferowane modele (na wykresie widnieje marka oraz konkretny model) na portalu OtoMoto zostały umieszczone w odpowiednio posortowanej liście na poniższej infografice:

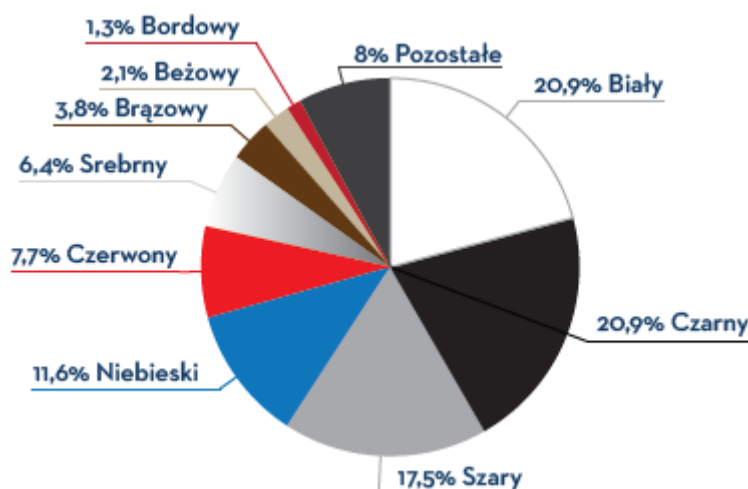


Rys 19. Najczęściej oferowane modele samochodów (kwiecień 2021)

Obserwując polskie drogi, nie powinny dziwić wyniki przedstawione powyżej. Wymienione zostały popularne marki – zarówno mniej oraz bardziej prestiżowe – oraz najmocniej kojarzone z nimi modele.

Wspomniana już statystyka jest możliwa do sprawdzenia w sposób empiryczny, po udaniu się na krótki spacer chodnikiem, usadowionym nieopodal ruchliwej ulicy w większym polskim mieście. Kolejną z takich kategorii jest kolor powłoki lakierniczej oferowanych pojazdów. W roku 2018, według informacji zamieszczonych w raporcie, najczęściej poszukiwanym kolorem były czarny oraz biały. Można zatem stwierdzić, że klasyka zawsze

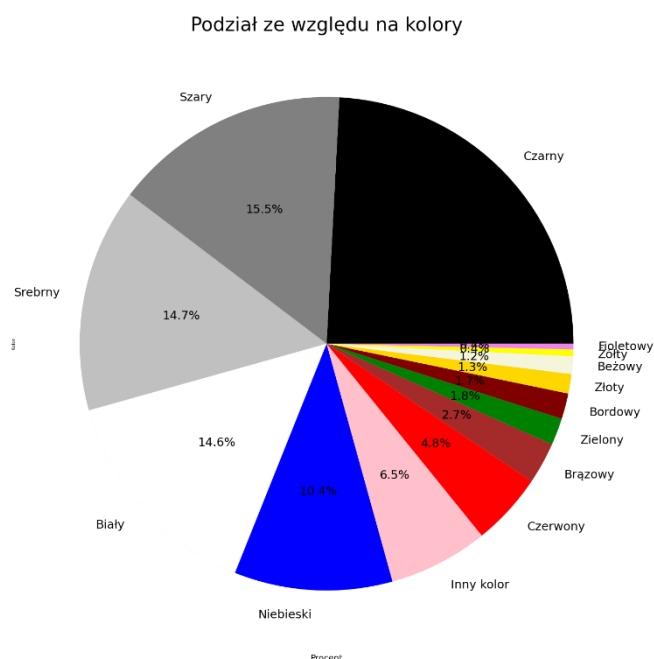
pozostaje w modzie. Na następnym miejscu plasuje się również – niezbyt ekstrawagancka barwa: szara. Pozostałe kolory lakieru, są zdecydowanie mniej popularne.



Rys 20. Rozkład procentowy wyszukiwań samochodów ze względu na kolor nadwozia (OtoMoto 2018)

Obecnie należałoby to skonfrontować z danymi pozyskanymi na potrzeby projektu, a więc z takimi, jakie możliwe są do odczytania bezpośrednio ze strony internetowej. Jak widać na załączonej poniżej grafice, pewne tendencje rynku motoryzacyjnego pozostają bez zmian – niezależnie czy mowa o poszukiwaniach, czy też dostępnych na rynku wtórnym. Występuje w tym przypadku już znacznie większa dysproporcja między najpopularniejszymi kolorami, ponieważ klasyczna czerń posiada znaczną przewagę nad zbliżonymi wartościami: szarym, srebrnym oraz białym.

Można zatem śmiało stwierdzić, ponownie – również z doświadczeń empirycznych – iż na polskich ulicach dominują popularnością obiekty motoryzacyjne, o niewysublimowanych kolorach – tych, które postrzegane są jako najbardziej klasyczne.



Rys 21. Rozkład procentowy dostępnych samochodów ze względu na kolor nadwozia (kwiecień 2021)

Po przeanalizowaniu najbardziej popularnych oraz typowych „konkurencji” obrazujących w jak największym stopniu rynek osobowych samochodów używanych oraz nowych, wyrobione zostało pewne postrzeganie rzeczywistości motoryzacyjnego świata w Polsce. Warto jednak czasami opuścić konwencjonalne i sztywnie trzymane w ramach myślenie i skupić się na marzeniach. Przeważnie najbardziej pożądanymi samochodami są te, które oferują najlepsze osiągi: potężne silniki, wspaniałe właściwości jezdne i niejednokrotnie absurdalnie wysokie ceny. Jednakowoż marzyć należy z rozmachem, zatem poniżej zaprezentowane zostały pojazdy osobowe z najmocniejszymi jednostkami napędowymi. W tabeli znajdują się informacje takie jak: id ogłoszenia, marka, model, moc silnika, cena wywoławcza pobrana z ogłoszenia oraz waluta odnosząca się do tej poprzedniej miary.

	Marka	Model	Moc	Cena	Waluta
46904	dodge	charger	885.0	310000.0	PLN
46971	dodge	charger	850.0	259000.0	PLN
47522	ferrari	488	800.0	1569000.0	PLN
47552	ferrari	488	800.0	738000.0	EUR
47562	ferrari	488	800.0	1299000.0	PLN
47585	ferrari	812-superfast	800.0	1569000.0	PLN
47586	ferrari	812-superfast	800.0	738000.0	EUR
47587	ferrari	812-superfast	800.0	1299000.0	PLN
106212	mercedes-benz	g-klasa	800.0	510000.0	EUR
187205	toyota	supra	799.0	550000.0	PLN

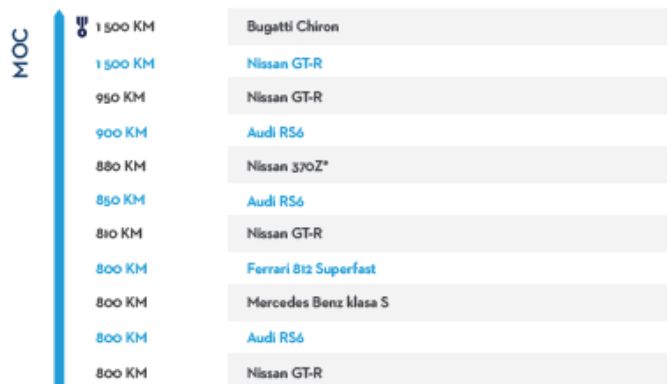
Rys 22. Dziesięć samochodów osobowych o najwyższej mocy silnika (kwiecień 2021)

Amerykańska marka zajęła aż dwa miejsca na podium listy dzięki swojemu modelowi: Charger, lecz ogólnie ów lista została zdominowana przez włoską motoryzację, ponieważ aż sześć spośród dziesięciu pozycji należy do marki Ferrari.



Rys 23. Dodge Charger legitymujący się mocą silnika: 885KM
(najmocniejszy samochód osobowy – kwiecień 2021)

Spoglądając jednak na zbliżoną statystykę pochodzącą z 2018 roku, powyższe statystyki nie prezentują się tak okazale, jak mogłoby się wydawać. Co oczywiste, różnica ta może wynikać z tego, iż w przypadku realizowanego projektu, pod uwagę brane były jedynie ogłoszenia będące aktywnymi w trakcie trwania procesu scrapingu. Analitycy serwisu OtoMoto posiadali naturalnie dane zbierane przez cały okres funkcjonowania serwisu, zatem ów dane są bardziej rzetelne w perspektywie całego roku.



*Według opisu wystawiającego, to jedyny taki Nissan 370z w Polsce. Najmocniejsza auto serii Z w naszym kraju, dysponujące mocą regulowaną jednym guzikiem (boost) 500/750/880 HP i 900 Nm.

Rys 24. Dziesięć samochodów osobowych o najwyższej mocy silnika (OtoMoto 2018)

Wiedząc już jakie samochody mogą dostarczyć najwięcej emocji w trakcie prowadzenia, należałoby skierować swój wzrok ku samochodom najbardziej luksusowym, prestiżowym. Poniżej przedstawione zostały najdroższe dostępne na aukcjach w kwietniu samochody oraz odpowiadające tej samej kategorii pojazdy z całego roku 2018 (10) oraz 2020 (5). W przypadku statystyk sporządzonych na potrzeby projektu studenckiego, wybór został ograniczony do pojazdów, których cena wyjściowa została podana w polskiej walucie.

	Marka	Model	Cena	Waluta	Stan
108232	mercedes-benz	slr	3490000.0	PLN	Używane
147316	porsche	911	2583000.0	PLN	Używane
159396	rolls-royce	ghost	2370000.0	PLN	Nowe
159394	rolls-royce	cullinan	2330000.0	PLN	Nowe
89842	lamborghini	aventador	2300000.0	PLN	Używane
159416	rolls-royce	wraith	2101970.0	PLN	Nowe
159393	rolls-royce	cullinan	2100000.0	PLN	Używane
159419	rolls-royce	wraith	2013000.0	PLN	Nowe
159417	rolls-royce	wraith	1950000.0	PLN	Nowe
97320	mclaren	720s-spider	1862200.0	PLN	Nowe

Rys 25. Dziesięć najdroższych samochodów osobowych (PLN) (kwiecień 2021)



Rys 26. Dziesięć najdroższych samochodów osobowych (OtoMoto 2018)

TOP 5
NAJDRÓŻSZYCH OFERT SPRZEDAŻY SAMOCHODÓW

MARKA	MODEL	ROK PRODUKCJI	CENA
BUGATTI	CHIRON	2020	17 700 000 zł
FERRARI	LA FERRARI	2014	9 417 000 zł
LAMBORGHINI	AVENTADOR SVJ	2020	5 250 000 zł
BUGATTI	VEYRON	2009	3 827 000 zł
MERCEDES BENZ	SLR	2009	3 490 000 zł

Rys 27. Pięć najdroższych samochodów osobowych (OtoMoto 2020)

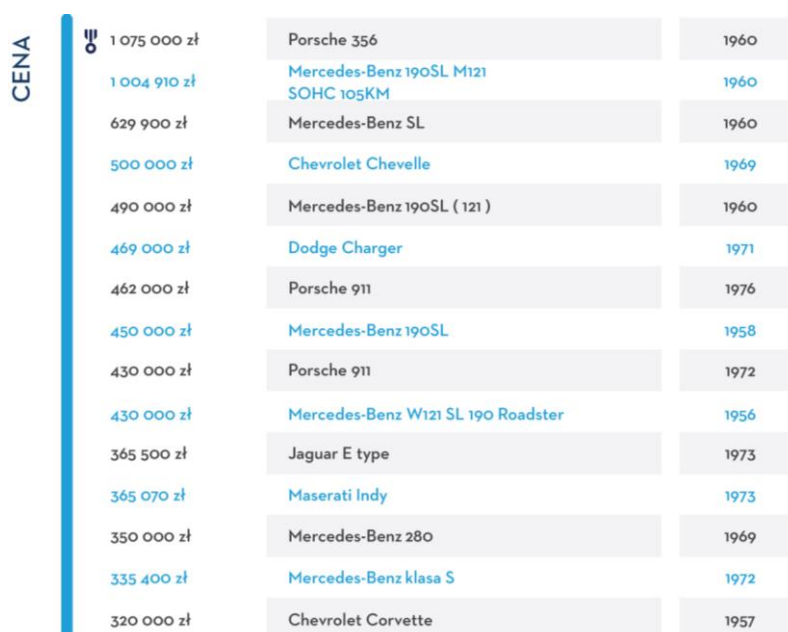
Zmierzając ku końcowi przedstawianych w projekcie efektów dokonanej analizy, sporządzona została jeszcze jedna lista, która powinna zostać zaklasyfikowana bliżej tych dotyczących motoryzacyjnych marzeń, niżli danych opowiadających o realnych trendach rynku osobowych samochodów używanych oraz nowych w Polsce.

Mowa o nostalgicznym powiewie, dojrzałego już wiatru, który przywodzi na myśl czasy zupełnie odmienne od tych obecnie panujących. Czasów, w których niemalże każdy obszar konstrukcyjny był wykonany w pełni w formie analogowej. Opis oczywiście dotyczy pojazdów zabytkowych. Tak jak miało to miejsce w przypadku wcześniejszych „konkurencji”, tak i tutaj zestawione zostaną wyniki analizy z kwietnia 2021 roku oraz te, które zawarte zostały w raporcie za cały rok 2018.

By dany pojazd mógł zostać uznany za zabytek, w świetle polskiego prawa, musi spełniać szereg warunków, takich jak między innymi: model ten nie może być obecnie produkowany (nie uwzględniając nowszych generacji), musi być starszy niż 30 lat, w przypadku samochodów seryjnie produkowanych oraz posiadać co najmniej 75% oryginalnych części.

	Marka	Model	Rok prod.	Cena	Waluta	Czy zabytek
80254	jaguar	xk	1956.0	890000.0	PLN	True
79406	jaguar	e-type	1965.0	725000.0	PLN	True
47544	ferrari	488	1986.0	685000.0	PLN	True
47615	ferrari	testarossa	1986.0	685000.0	PLN	True
147275	porsche	911	1976.0	650000.0	PLN	True
66127	ford	mustang	1967.0	499000.0	PLN	True
66534	ford	mustang	1968.0	498000.0	PLN	True
66265	ford	mustang	1967.0	480000.0	PLN	True
147354	porsche	911	1968.0	455000.0	PLN	True
47500	ferrari	308	1978.0	390000.0	PLN	True

Rys 28. Dziesięć najdroższych samochodów osobowych, zabytkowych (kwiecień 2021)



Rys 29. Dziesięć najdroższych samochodów osobowych, zabytkowych (OtoMoto 2018)

Ponownie uwidocznione zostały różnice pomiędzy poszczególnymi latami dokonywania analizy. Tak jak miało to miejsce w przypadku wcześniejszych porównań, raport utworzony przez OtoMoto obejmuje analizę całego roku kalendarzowego, w trakcie którego – co pewne – przewinęło się przez portal znacznie więcej interesujących klasyków, zabytkowych pojazdów.

W obrębie statystyk sporządzanych na potrzeby projektu studenckiego, uwzględnione zostały jedynie samochody, których cena wyjściowa aukcji, podana była w złotych. Prym

w tej kategorii (wszakże miejsca: pierwsze oraz drugie) wiedzy motoryzacja brytyjska. Zdecydowanie najdroższym jednak egzemplarzem jest XK i pochodzi z 1956 roku. Zdjęcie przedstawiające ów zabytek możliwe jest do wglądu poniżej:



Rys 26. Jaguar XK z 1956 roku (najdroższy samochód osobowy, zabytkowy – kwiecień 2021)

6. Bibliografia

- Repozytorium projektu: <https://github.com/nadrowskyy/otomoto-scrap-analyze>
- OtoMoto: Internetowy Samochód Roku
 - 2018 (https://www.otomoto.pl/isr/otomoto_raport2018.pdf)
 - 2020 (<https://motopedia.otomoto.pl/katalog-isr-2020>)
- <https://cloud.google.com/docs>
- <https://pandas.pydata.org/docs/>
- <https://autokult.pl/34121,koniec-z-zabytkowa-patologia-ministerstwo-kultury-ustali-nowe-warunki-dla-zoltych-tablic>