Analiza konkretnego zbioru danych dotyczącego wypożyczania rowerów może prowadzić do wielu interesujących celów badawczych i wniosków. Poniżej przedstawiam kilka potencjalnych celów analizy tego typu danych:

1. **Prognozowanie popytu**: Można wykorzystać dane historyczne dotyczące wypożyczeń rowerów do prognozowania przyszłego popytu na wypożyczenia w różnych warunkach, takich jak pogoda, dzień tygodnia, pora roku itp. To może być przydatne dla firm wypożyczalni rowerów do planowania zasobów i optymalizacji usług.
2. **Segmentacja użytkowników**: Analiza danych może pomóc w identyfikacji różnych grup użytkowników wypożyczalni rowerów na podstawie ich zachowań, preferencji i cech demograficznych. Pozwala to na lepsze zrozumienie potrzeb różnych segmentów klientów i dostosowanie oferty do ich potrzeb.
3. **Optymalizacja lokalizacji stacji wypożyczalni**: Przeanalizowanie danych dotyczących lokalizacji stacji wypożyczalni i wzorców wypożyczeń może pomóc w optymalnym rozmieszczeniu nowych stacji w celu zwiększenia dostępności usług i zaspokojenia popytu.
4. **Analiza wydajności stacji**: Można ocenić wydajność poszczególnych stacji wypożyczalni pod kątem liczby wypożyczeń, zwrotów, popularności itp. To może prowadzić do lepszego zarządzania zasobami i optymalizacji działania stacji.
5. **Wpływ czynników zewnętrznych na wypożyczenia**: Badanie, jak czynniki zewnętrzne, takie jak pogoda, święta, wydarzenia specjalne, wpływają na liczność wypożyczeń, może pomóc w planowaniu działań marketingowych i promocyjnych.
6. **Analiza trendów czasowych**: Analiza trendów czasowych w danych może pomóc w identyfikacji sezonowych wzorców wypożyczeń oraz długoterminowych trendów, co może prowadzić do lepszego planowania i strategii rozwoju biznesu.
7. **Ocena wpływu promocji i rabatów**: Badanie wpływu różnych promocji, rabatów i ofert specjalnych na liczbę wypożyczeń może dostarczyć informacji zwrotnych na temat skuteczności działań marketingowych.
8. **Analiza zachowań użytkowników**: Przeanalizowanie danych dotyczących zachowań użytkowników, takich jak długość wypożyczeń, trasy podróży, częstotliwość korzystania z usługi, może pomóc w zrozumieniu preferencji i potrzeb klientów oraz dostosowaniu oferty do nich.
9. **Ocena jakości usług**: Analiza opinii i ocen użytkowników może dostarczyć informacji zwrotnych na temat jakości usług i pomóc w identyfikacji obszarów do poprawy.
10. **Analiza opłacalności**: Przeanalizowanie danych finansowych w połączeniu z danymi dotyczącymi wypożyczeń może pomóc w ocenie opłacalności biznesu wypożyczalni rowerów oraz identyfikacji obszarów do optymalizacji kosztów lub zwiększenia przychodów.

Te cele stanowią jedynie przykłady możliwych kierunków analizy danych dotyczących wypożyczania rowerów. Ostateczny wybór celu zależy od konkretnych potrzeb i celów biznesowych organizacji zbierającej i analizującej te dane.

 **Analiza sezonowa**:

* Sprawdź, jak liczba wypożyczeń rowerów (kolumna cnt) zmienia się w zależności od pory roku (season).
* Wykorzystaj formułę SUMIF, np. =SUMIF(season, 1, cnt) dla wiosny.
* **Wykres**: Wykres słupkowy przedstawiający sumaryczną liczbę wypożyczeń dla każdej pory roku.

 **Analiza roczna**:

* Porównaj liczbę wypożyczeń rowerów w latach 2011 i 2012 (yr).
* Użyj formuły SUMIF, np. =SUMIF(yr, 0, cnt) dla 2011 roku.
* **Wykres**: Wykres słupkowy porównujący sumaryczną liczbę wypożyczeń w latach 2011 i 2012.

 **Wpływ dni roboczych i świąt**:

* Zbadaj różnice w liczbie wypożyczeń między dniami roboczymi (workingday) a weekendami/świętami.
* Użyj formuły SUMIF, np. =SUMIF(workingday, 1, cnt) dla dni roboczych.
* **Wykres**: Wykres słupkowy porównujący liczbę wypożyczeń w dni robocze i wolne od pracy.

 **Wpływ pogody**:

* Sprawdź, jak różne warunki pogodowe (weathersit) wpływają na liczbę wypożyczeń.
* Użyj formuły SUMIF, np. =SUMIF(weathersit, 1, cnt) dla dni z pogodą "Clear, Few clouds, Partly cloudy".
* **Wykres**: Wykres słupkowy pokazujący liczbę wypożyczeń dla różnych warunków pogodowych.

 **Analiza miesięczna**:

* Zbadaj, jak liczba wypożyczeń rowerów zmienia się w ciągu roku (kolumna mnth).
* Użyj formuły SUMIF, np. =SUMIF(mnth, 1, cnt) dla stycznia.
* **Wykres**: Wykres liniowy pokazujący liczbę wypożyczeń w poszczególnych miesiącach.

 **Wpływ wilgotności i prędkości wiatru**:

* Zbadaj, jak wilgotność (hum) i prędkość wiatru (windspeed) wpływają na liczbę wypożyczeń.
* Użyj formuły CORREL, np. =CORREL(hum, cnt) dla korelacji między wilgotnością a liczbą wypożyczeń.
* **Wykres**: Dwa wykresy punktowe, jeden pokazujący zależność między wilgotnością a liczbą wypożyczeń, a drugi między prędkością wiatru a liczbą wypożyczeń.

 **Analiza według dni tygodnia**:

* Sprawdź, jak liczba wypożyczeń zmienia się w zależności od dnia tygodnia (weekday).
* Użyj formuły SUMIF, np. =SUMIF(weekday, 1, cnt) dla poniedziałków.
* **Wykres**: Wykres słupkowy przedstawiający liczbę wypożyczeń dla poszczególnych dni tygodnia.

 **Porównanie użytkowników zarejestrowanych i okazjonalnych**:

* Porównaj liczbę wypożyczeń przez zarejestrowanych (registered) i okazjonalnych użytkowników (casual).
* Użyj formuły SUM, np. =SUM(registered) i =SUM(casual) dla sumarycznej liczby wypożyczeń w obu grupach.
* **Wykres**: Wykres słupkowy porównujący liczbę wypożyczeń przez zarejestrowanych i okazjonalnych użytkowników.

 **Wpływ świąt na wypożyczenia**:

* Zbadaj, czy liczba wypożyczeń różni się w dni świąteczne (holiday).
* Użyj formuły SUMIF, np. =SUMIF(holiday, 1, cnt) dla dni świątecznych.
* **Wykres**: Wykres słupkowy porównujący liczbę wypożyczeń w dni świąteczne i zwykłe dni.

ML

Analiza danych z wykorzystaniem Machine Learning na przykładzie danych o wynajmowaniu rowerów składa się z kilku kluczowych etapów, a każdy z nich prowadzi do istotnych wniosków.

1. **Wczytanie danych**: Dane są wczytywane z pliku CSV do programu. Dane dotyczą wynajmu rowerów i zawierają informacje o różnych cechach, takich jak temperatura, wilgotność, dzień tygodnia itp.
2. **Przygotowanie danych**: Dane są przygotowywane do analizy. Skalowane są wartości kolumn, takich jak temperatura, aby były one w odpowiednich jednostkach. Następnie usuwane są kolumny, które nie są potrzebne do analizy, na przykład kolumna z datą.
3. **Modelowanie**: Regresja liniowa jest stosowana, aby określić, jak różne cechy wpływają na liczbę wynajętych rowerów. Model jest trenowany na całym zestawie danych.
4. **Wyniki modelu**: Po trenowaniu modelu analizowane są współczynniki regresji. Wartości te wskazują, jak silnie każda cecha wpływa na liczbę wynajętych rowerów. Na przykład wysoki współczynnik dla temperatury oznacza, że temperatura ma duży wpływ na wynajem rowerów.
5. **Wizualizacja wyników**: Tworzone są wykresy w celu ułatwienia interpretacji wyników. Macierz korelacji pokazuje, jak różne cechy są ze sobą powiązane. Heatmapa (mapa cieplna) wizualizuje te zależności, co pozwala zobaczyć, które cechy są ze sobą silnie powiązane, a które mniej.
6. **Analiza wykresów**:
   * **Macierz korelacji**: Pozwala zobaczyć, jak różne czynniki wpływają na siebie nawzajem. Na przykład można zauważyć, że liczba wynajętych rowerów jest silnie powiązana z temperaturą.
   * **Wykresy punktowe**: Pokazują relację między konkretną cechą (np. temperatura) a liczbą wynajętych rowerów. Można zauważyć, że w cieplejsze dni rowery są częściej wynajmowane.

Podsumowując, taka analiza umożliwia zrozumienie, jakie czynniki wpływają na liczbę wynajmowanych rowerów. Na przykład można się dowiedzieć, że temperatura i dzień tygodnia mają duży wpływ, a święta mogą negatywnie wpływać na wynajem. Dzięki takim analizom można lepiej planować działania, na przykład zwiększać dostępność rowerów w cieplejsze dni lub przygotowywać specjalne oferty na dni świąteczne, aby zachęcić ludzi do wynajmu.