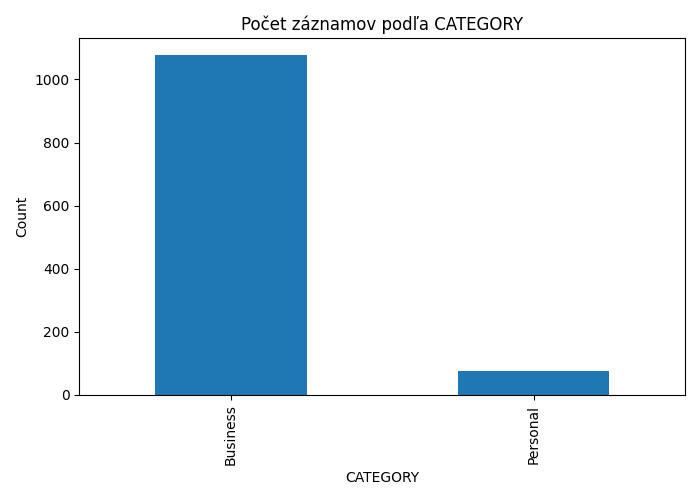
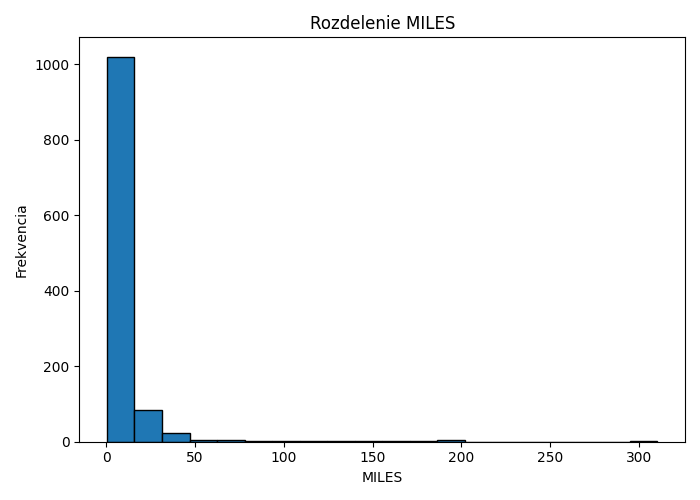
Uber Dataset(2016)

## Analýza grafov a model vs. Baseline

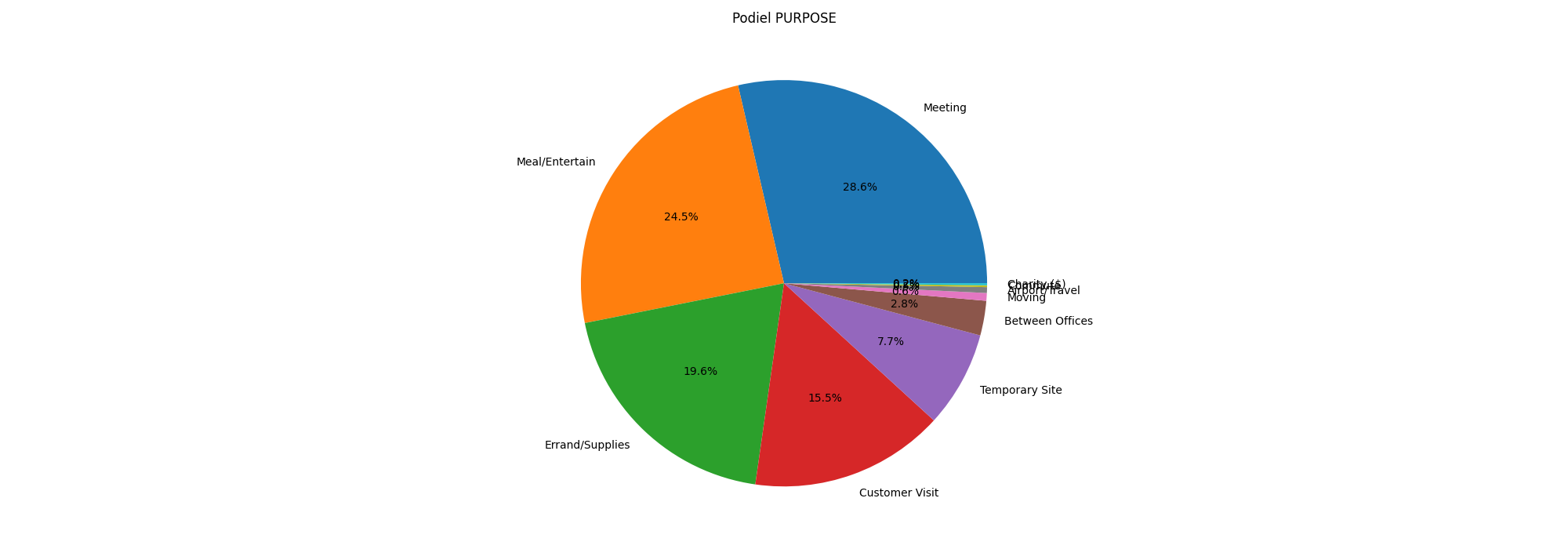
Ako môžeme vidieť, Uber sa primárne používa na business jazdy (93,33 %) a s malou odchýlkou na personal jazdy (6,66 %).



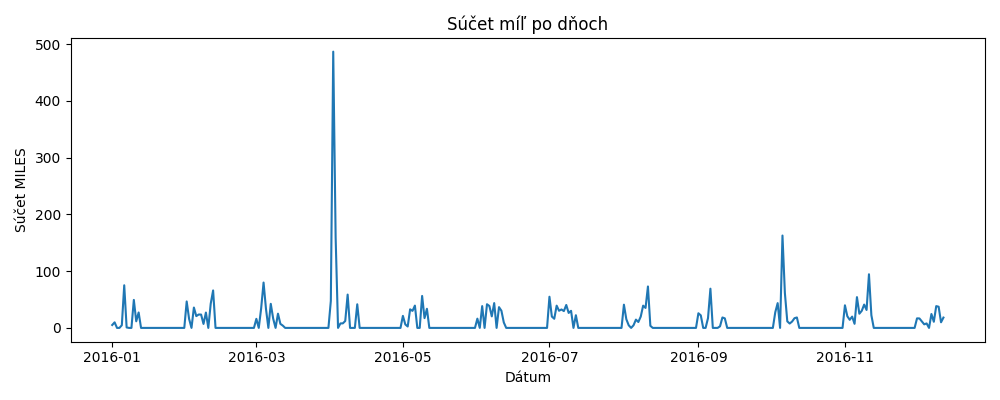
Druhý graf nám ukazuje, že najväčší počet jázd je iba do 15 míľ (1 míľa ≈ 1,6 km).



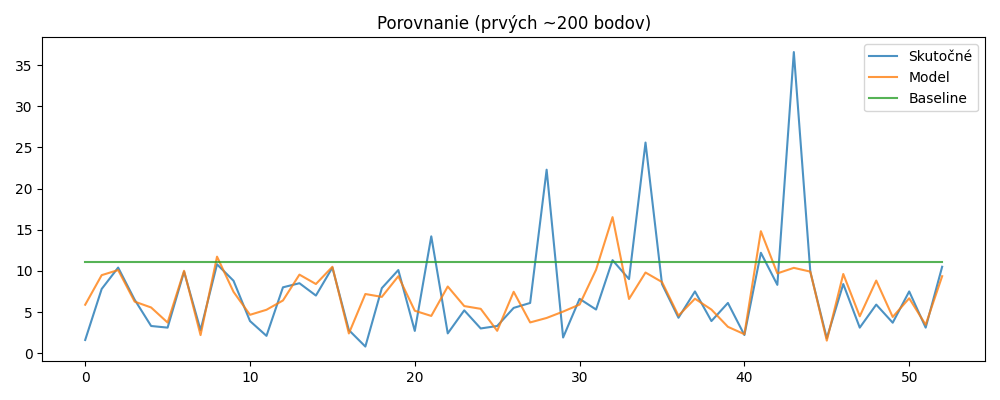
A koláčový graf nám ukazuje, že štyri najčastejšie ciele boli: stretnutie s niekým (28,6 %), jedlo/zábava (24,5 %), nákupy (19,6 %) a návšteva zákazníkov (15,5 %).



Ďalší obrázok nám ukazuje, že najviac míľ bolo prejdených približne v apríli a druhý pik bol v októbri.



MAE: ukazuje, ako veľmi sa model zvyčajne mýli (priemer absolútnej hodnoty rozdielov medzi výsledkom modelu a reálneho výsledku, hodnota má byť čo najnižšia).  
MSE: ukazuje chyby a viac zvýrazňuje veľké omyly (každá chyba má svoju váhu, priemer druhej mocniny čísel z vyššej vysvetlivky).  
R²: hovorí, či model dobre vysvetľuje dáta alebo nie (ako často model lepšie trafi realitu ako baseline, ak výsledok je medzi 1 a 0 tak je to najviac priblížené ku realite, ak je záporný tak je to veľmi zlý model na predhadanie výsledkov).  
Baseline je to iba priemer, ktorý nám nedokáže ukázať plný obrázok reality. A ako vidíme, model je o veľa presnejší: má iba MAE: 2.825, MSE: 30.208, R²: 0.207, keď baseline má MAE: 5.718, MSE: 51.095, R²: -0.342.



Také modely sa naučia na reálnych dátach a potom môzhu predpovedať budúcnosť, ale čím ďalej od súčasného momentu, tým nižšia je presnosť predpovede.