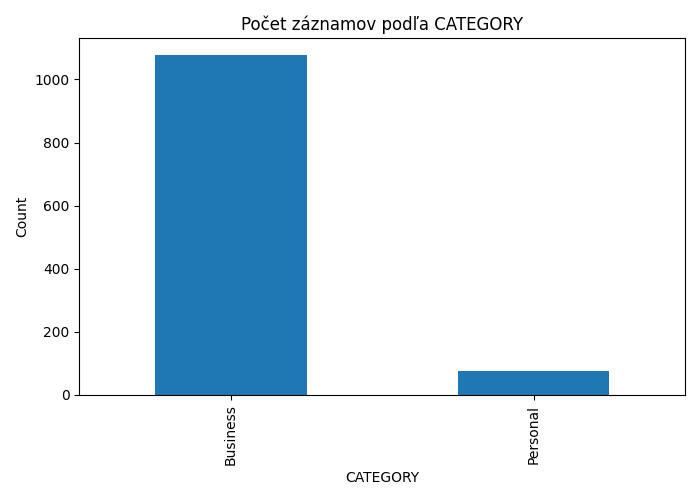
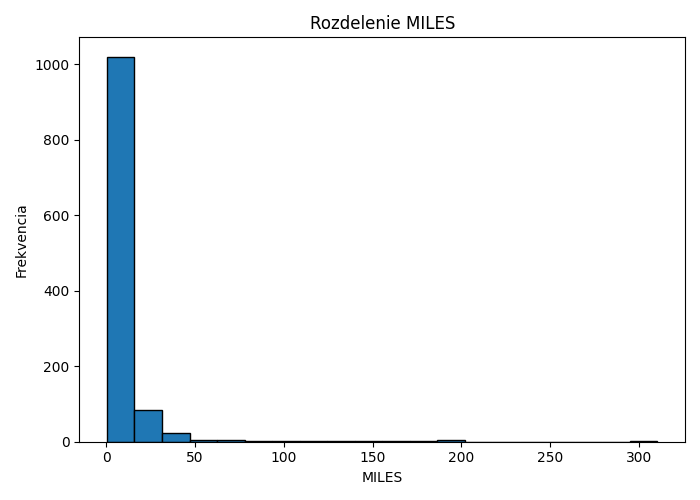
Uber Dataset(2016)

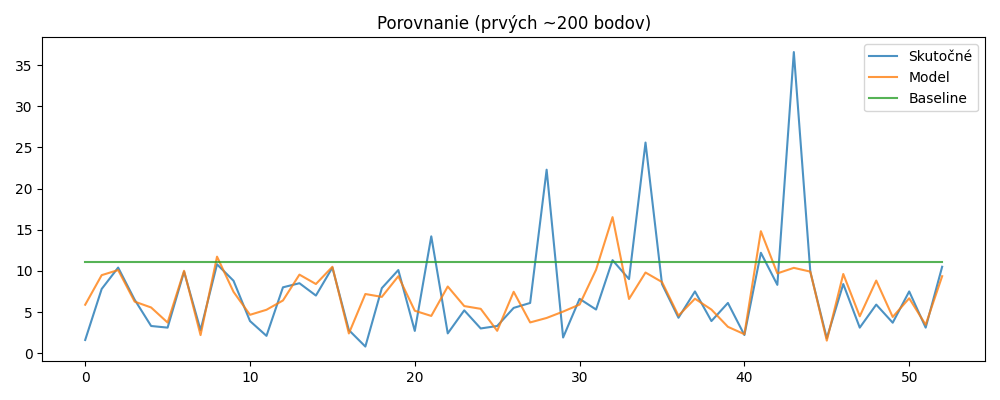
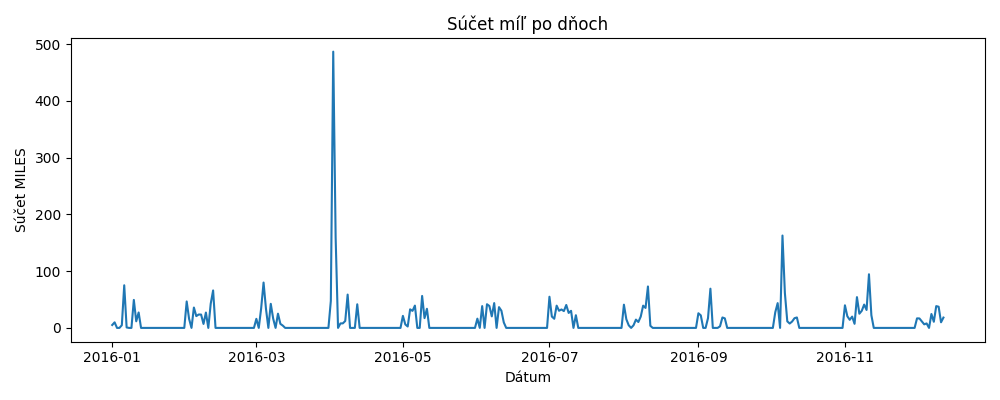
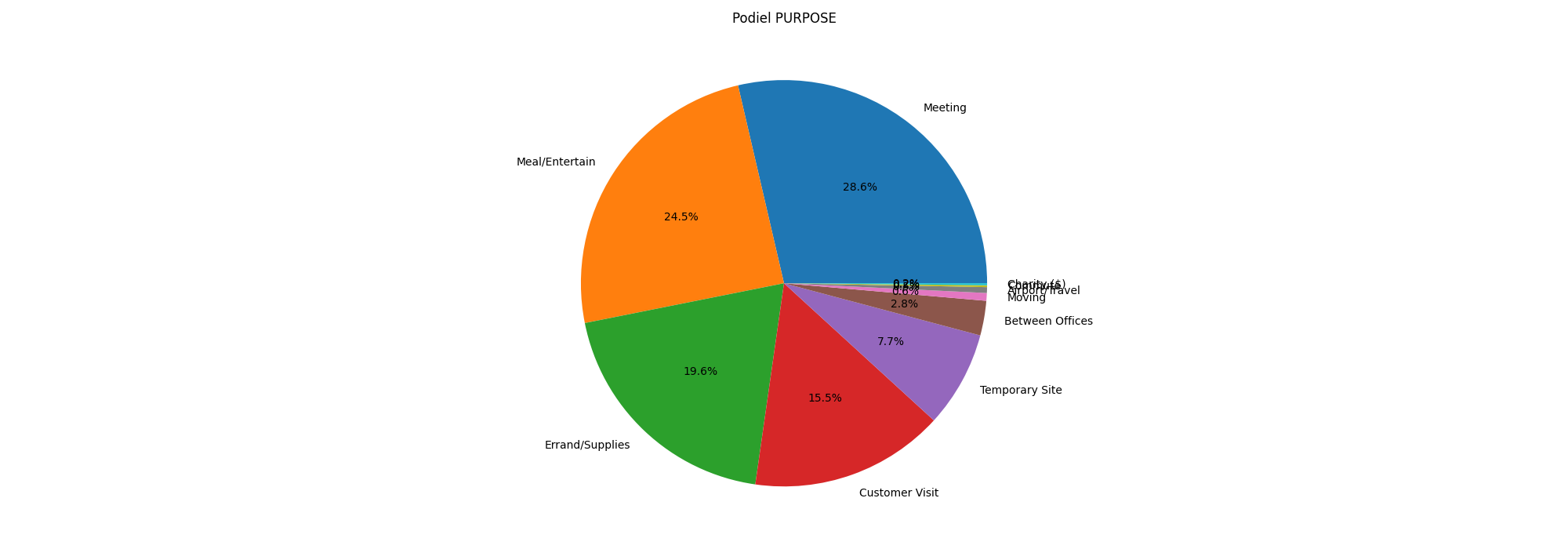
## analiza grafov a model vs. Baseline



Ako mozeme vidiet Uber sa primarne pouziva na Business jazdy(93.33%) s malou odchylkou na Personal(6.66%).



Druhy graf nam ukazuje, ze najvacsi pocet jaz je do (1 mila ~1.6 km)



Baseline je to iba priemer, ktorý nám nedokáže ukázať

--- Hlava dát ---

START\_DATE END\_DATE CATEGORY START STOP MILES PURPOSE

0 01-01-2016 21:11 01-01-2016 21:17 Business Fort Pierce Fort Pierce 5.1 Meal/Entertain

1 01-02-2016 01:25 01-02-2016 01:37 Business Fort Pierce Fort Pierce 5.0 NaN

2 01-02-2016 20:25 01-02-2016 20:38 Business Fort Pierce Fort Pierce 4.8 Errand/Supplies

3 01-05-2016 17:31 01-05-2016 17:45 Business Fort Pierce Fort Pierce 4.7 Meeting

4 01-06-2016 14:42 01-06-2016 15:49 Business Fort Pierce West Palm Beach 63.7 Customer Visit

--- Info ---

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 1156 entries, 0 to 1155

Data columns (total 7 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 START\_DATE 1156 non-null object

1 END\_DATE 1155 non-null object

2 CATEGORY 1155 non-null object

3 START 1155 non-null object

4 STOP 1155 non-null object

5 MILES 1156 non-null float64

6 PURPOSE 653 non-null object

dtypes: float64(1), object(6)

memory usage: 63.3+ KB

None

--- Popisné štatistiky ---

START\_DATE END\_DATE CATEGORY START STOP MILES PURPOSE

count 1156 1155 1155 1155 1155 1156.000000 653

unique 1155 1154 2 177 188 NaN 10

top 6/28/2016 23:34 6/28/2016 23:59 Business Cary Cary NaN Meeting

freq 2 2 1078 201 203 NaN 187

mean NaN NaN NaN NaN NaN 21.115398 NaN

std NaN NaN NaN NaN NaN 359.299007 NaN

min NaN NaN NaN NaN NaN 0.500000 NaN

25% NaN NaN NaN NaN NaN 2.900000 NaN

50% NaN NaN NaN NaN NaN 6.000000 NaN

75% NaN NaN NaN NaN NaN 10.400000 NaN

max NaN NaN NaN NaN NaN 12204.700000 NaN

Počet chýbajúcich PURPOSE: 502

Train shape: (208, 150), Test shape: (53, 150)

=== Výsledky modelu ===

MAE: 2.825

MSE: 30.208

R² : 0.207

=== Baseline (priemer tréningu) ===

MAE: 5.718

MSE: 51.095

R² : -0.342

Hotovo. Grafy nájdeš v priečinku 'figures/'.