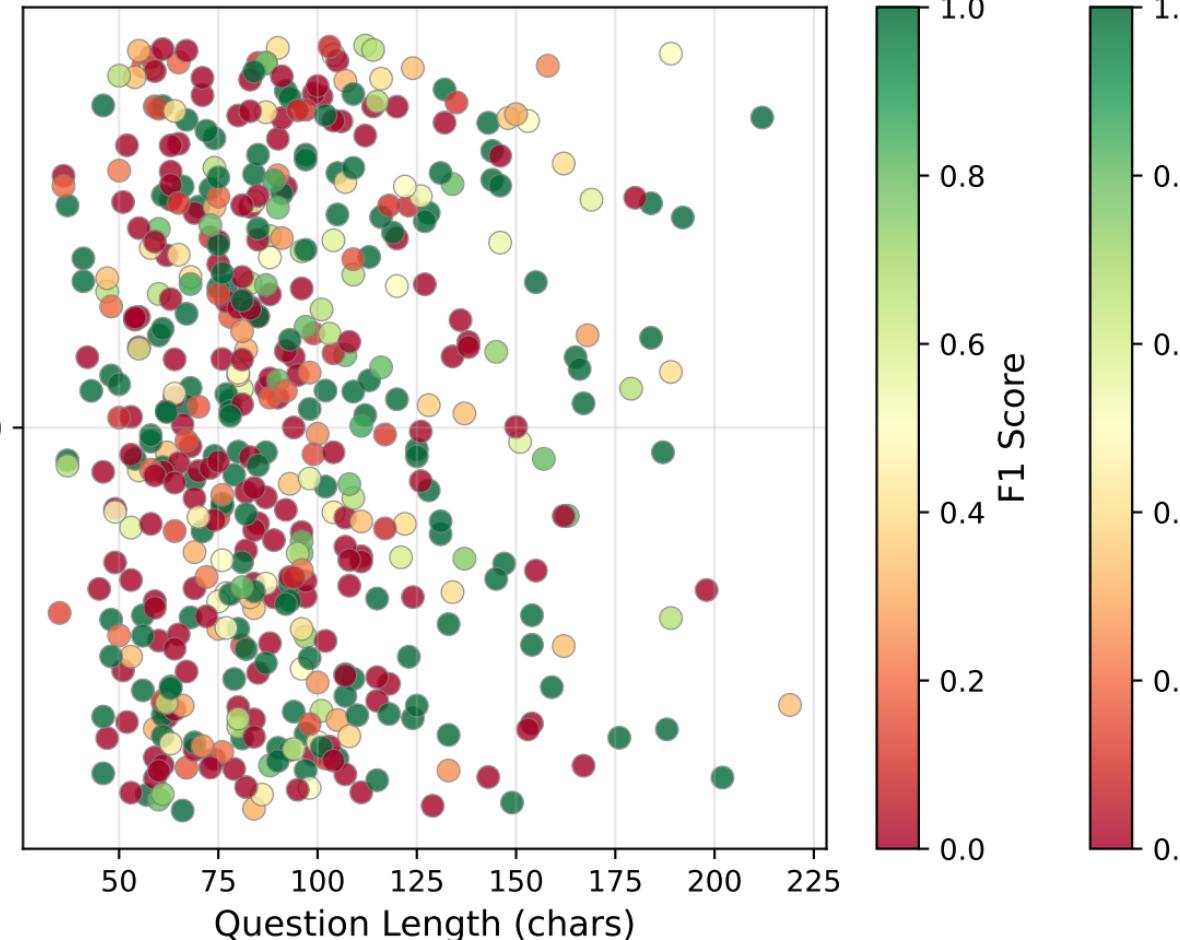


TopK Decision vs Question Length

Chosen TopK

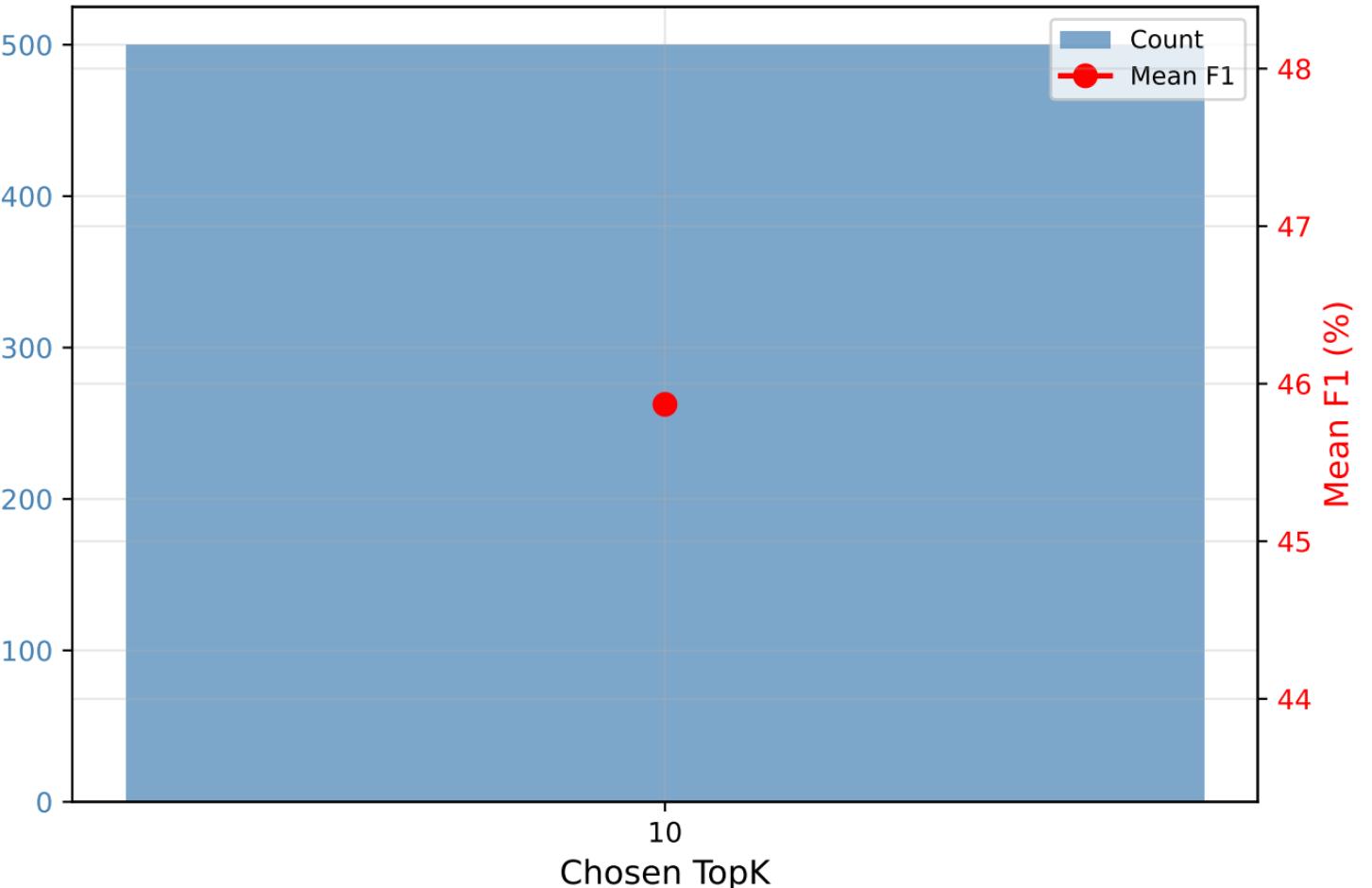


F1 Score

1.0
0.8
0.6
0.4
0.2
0.0

TopK Distribution and F1 by Bucket

Number of Queries



Mean F1 (%)

48
47
46
45
44

Question Length (chars)

Chosen TopK