Common Ideas in State of the Art Retrieval Models

f(q="presidential campaign news", d)

"Bag of Words"

g("presidential", d)   g("campaign", d)   g("news", d)

How many times does "presidential" occur in d?
Term Frequency (TF):   c("presidential", d)

How long is d?   **Document length**: |d|

How often do we see "presidential" in the entire collection?
Document Frequency: df("presidential")
P("presidential"|collection)