MongoDB Twitter Exercise

1. **How many Twitter users are in the database?**

db. tweets.distinct(“user”).size();

1. **Which Twitter users link the most to other Twitter users? (Provide the top ten.)**

db.tweets.aggregate(

{$match: {text: {$regex: "(?<=^|(?<=[^a-zA-Z0-9-\_\\\\.]))@([A-Za-z]+[A-Za-z0-9\_]+)" }}},

{$group: {\_id: $user, “count”: {$sum: 1}}},

{$sort: {“count”: -1}},

{$limit: 10}

);

1. **Who is are the most mentioned Twitter users? (Provide the top five.)**

db.tweets.aggregate(

{$match: {text: {$regex: "(?<=^|(?<=[^a-zA-Z0-9-\_\\\\.]))@([A-Za-z]+[A-Za-z0-9\_]+)”}}},

{ … ?},

{$group: {\_id: $user, “count”: {$sum: 1}}},

{$sort: {“count”: -1}},

{$limit: 5}

);

1. **Who are the most active Twitter users (top ten)?**

db.tweets.aggregate(

{$group: {\_id: $user, “count”: {$sum: 1}}},

{$sort: {“count”: -1}},

{$limit: 10}

);

1. **Who are the five most grumpy (most negative tweets) and the most happy (most positive tweets)? (Provide five users for each group)**

db.tweets.aggregate(

{$match: {polarity: 0}},

{$group: {\_id: $user, “count”: {$sum: 1}}},

{$sort: {“count”: -1}},

{$limit: 5}

);

db.tweets.aggregate(

{$match: {polarity: 4}},

{$group: {\_id: $user, “count”: {$sum: 1}}},

{$sort: {“count”: -1}},

{$limit: 5}

);

**Links:**

<https://docs.mongodb.com/manual/reference/operator/aggregation/match/>

<https://docs.mongodb.com/manual/reference/operator/aggregation/group/>

<https://docs.mongodb.com/manual/reference/operator/aggregation/sort/>

<https://docs.mongodb.com/manual/reference/operator/aggregation/limit/>

<https://docs.mongodb.com/manual/reference/operator/aggregation/sum/#grp._S_sum>