Khushboo Tekchandani

CSE587 Homework 5

Problem:

This assignment required us to analyze the popularity of NBA teams using data from Twitter. We use the Twitter API to fetch recent tweets about the NBA teams in East and West Divisions. Based on occurrence of the words 'win' and 'lose', we determine the popularity of these teams.

Solution:

Step1 : Retrieve data

I used a python script to gather the most recent 100 tweets about each NBA team by filtering tweets that involved the Hashtags for each team.

Step2 : MapReduce

Using this data, I ran a simple map reduce job (similar to wordcount). The mapper takes the input from the file and maps each filename, which is a team's hashtag, with the occurrence of the word 'win' or 'lose'. To do this, the mapper writes out the <key,value> as <hashtag/win, 1> or <hashtag/lose,1>.

The reducer takes the input from the mapper. For each key, it increments a counter. Thus, I can count the number of wins or losses of a particular team. The output from the reducer is of the form (Row, Win/Lose, TeamName, Coast, Value) in two different tables for Wins and Losses.

I used Accumulo to create two different users 'east' and 'west', in a way that the data of east division NBA teams can be accessed only by east user and that of west division NBA teams can be accessed only by the west user. I fed the mapreduce output to two different tables 'Wins' and 'Losses' in accumulo. The visibility of the rows was fixed by the information 'coast' that was given in the reducer output.

Results:

	East	
Team Name	**Hashtag**	**Win**
Boston Celtics	#Celtics	0
New York Knicks	#Knicks	1
Philadelphia 76ers	#76ers	0
New Jersey Nets	#Nets	0
Toronto Raptors	#Raptors	0
Chicago Bull	#Bulls	0
Indiana Pacers	#Pacers	0
Milwaukee Bucks	#Bucks	0
Detroit Pistons	#Pistons	0
Cleveland Cavaliers	#Cavs	4
Miami Heat	#MiamiHeat	0
Orlando Magic	#OrlandoMagic	0
Atlanta Hawks	#Hawks	4
Charlotte Bobcats	#Bobcats	0
Washington Wizards	#Wizards	0

	East	
Team Name	**Hashtag**	**Lose**
Boston Celtics	#Celtics	0
New York Knicks	#Knicks	0
Philadelphia 76ers	#76ers	0
New Jersey Nets	#Nets	0
Toronto Raptors	#Raptors	0
Chicago Bull	#Bulls	4
Indiana Pacers	#Pacers	2
Milwaukee Bucks	#Bucks	0
Detroit Pistons	#Pistons	1
Cleveland Cavaliers	#Cavs	0
Miami Heat	#MiamiHeat	0
Orlando Magic	#OrlandoMagic	0
Atlanta Hawks	#Hawks	2
Charlotte Bobcats	#Bobcats	0
Washington Wizards	#Wizards	0

	West	
Team Name	**Hashtag**	**Win**
Oklahoma City	#okcthunder	0
Denver Nuggets	#Nuggets	0
Portland Trailblazers	#TrailBlazers	0
Utah Jazz	#UtahJazz	2
Minnesotta Timberwolves	#TWolves	0
LA Lakers	#Lakers	1
Phoenix Suns	#Suns	2
Golden State Warriors	#GSWarriors	20
L.A. Clippers	#Clippers	6
Sacramento Kings	#NBAKings	1
San Antonio Spurs	#GoSpursGo	1
Dallas Mavericks	#Mavs	1
New Orleans Hornets	#Hornets	0
Memphis Grizzlies	#Grizzlies	0
Houston Rockets	#Rockets	0

	West	
Team Name	**Hashtag**	**Lose**
Oklahoma City	#okcthunder	0
Denver Nuggets	#Nuggets	1
Portland Trailblazers	#TrailBlazers	3
Utah Jazz	#UtahJazz	0
Minnesotta Timberwolves	#TWolves	0
LA Lakers	#Lakers	1
Phoenix Suns	#Suns	0
Golden State Warriors	#GSWarriors	0
L.A. Clippers	#Clippers	1
Sacramento Kings	#NBAKings	0
San Antonio Spurs	#GoSpursGo	0
Dallas Mavericks	#Mavs	0
New Orleans Hornets	#Hornets	0
Memphis Grizzlies	#Grizzlies	1
Houston Rockets	#Rockets	0