PDSP 2024, Lecture 02, 8 August 2024

Set up the table as a list

- Assign a value to a variable: `variable = value`
- Each entry is a row
- Each row is a tuple of columns
- `(Venue, Team 1, Team 2, Toss winner, Match winner, Run target)`
- `Run target` is an integer, all other columns are text (`String`)

In [1]:
```
matchlist = [
    ("Chennai","RCB","CSK","RCB","CSK",174),
    ("Mohali","DC","PK","PK","PK",175),
    ("Kolkata","KKR","SRH","SRH","KKR",209),
    ("Jaipur","RR","LSG","RR","RR",194),
    ("Ahmedabad","GT","MI","MI","GT",169),
    ("Bengaluru","PK","RCB","RCB","RCB",177),
    ("Chennai","CSK","GT","GT","CSK",207),
    ("Hyderabad","SRH","MI","MI","SRH",278),
    ("Jaipur","RR","DC","DC","RR",186),
    ("Bengaluru","RCB","KKR","KKR","KKR",183),
    ("Lucknow","LSG","PK","LSG","LSG",200),
    ("Ahmedabad","SRH","GT","SRH","GT",163),
    ("Visakhapatnam","DC","CSK","DC","DC",192),
    ("Mumbai","MI","RR","RR","RR",126),
    ("Bengaluru","LSG","RCB","RCB","LSG",182),
    ("Visakhapatnam","KKR","DC","KKR","KKR",273),
    ("Ahmedabad","GT","PK","PK","PK",200),
    ("Hyderabad","CSK","SRH","SRH","SRH",166),
    ("Jaipur","RCB","RR","RR","RR",184),
    ("Mumbai","MI","DC","DC","MI",235),
    ("Lucknow","LSG","GT","LSG","LSG",164),
    ("Chennai","KKR","CSK","CSK","CSK",138),
    ("Mohali","SRH","PK","PK","SRH",183),
    ("Jaipur","RR","GT","GT","GT",197),
    ("Mumbai","RCB","MI","MI","MI",197),
    ("Lucknow","LSG","DC","LSG","DC",168),
    ("Mohali","PK","RR","RR","RR",148),
    ("Kolkata","LSG","KKR","KKR","KKR",162),
    ("Mumbai","CSK","MI","MI","CSK",207),
    ("Bengaluru","SRH","RCB","RCB","SRH",288),
    ("Kolkata","KKR","RR","RR","RR",224),
    ("Ahmedabad","GT","DC","DC","DC",90),
    ("Mohali","MI","PK","PK","MI",193),
    ("Lucknow","CSK","LSG","LSG","LSG",177),
    ("Delhi","SRH","DC","DC","SRH",267),
    ("Kolkata","KKR","RCB","RCB","KKR",223),
    ("Mohali","PK","GT","PK","GT",143),
    ("Jaipur","MI","RR","MI","RR",180),
    ("Chennai","CSK","LSG","LSG","LSG",211),
    ("Delhi","DC","GT","GT","DC",225),
    ("Hyderabad","RCB","SRH","RCB","RCB",207),
    ("Kolkata","KKR","PK","PK","PK",262),
    ("Delhi","DC","MI","MI","DC",258),
    ("Lucknow","LSG","RR","RR","RR",197),
    ("Ahmedabad","GT","RCB","RCB","RCB",201),
    ("Chennai","CSK","SRH","SRH","CSK",213),
    ("Kolkata","DC","KKR","DC","KKR",154),
    ("Lucknow","MI","LSG","LSG","LSG",145),
    ("Chennai","CSK","PK","PK","PK",163),
```

```
    ("Hyderabad","SRH","RR","SRH","SRH",202),
    ("Mumbai","KKR","MI","MI","KKR",170),
    ("Bengaluru","GT","RCB","RCB","RCB",148),
    ("Dharamsala","CSK","PK","PK","CSK",168),
    ("Lucknow","KKR","LSG","LSG","KKR",236),
    ("Mumbai","SRH","MI","MI","MI",174),
    ("Delhi","DC","RR","RR","DC",222),
    ("Hyderabad","LSG","SRH","LSG","SRH",166),
    ("Dharamsala","RCB","PK","PK","RCB",242),
    ("Ahmedabad","GT","CSK","CSK","GT",232),
    ("Kolkata","KKR","MI","MI","KKR",158),
    ("Chennai","RR","CSK","RR","CSK",142),
    ("Bengaluru","RCB","DC","DC","RCB",188),
    ("Delhi","DC","LSG","LSG","DC",209),
    ("Guwahati","RR","PK","RR","PK",145),
    ("Mumbai","LSG","MI","MI","LSG",215),
    ("Bengaluru","RCB","CSK","CSK","RCB",219),
    ("Hyderabad","PK","SRH","PK","SRH",215),
    ("Ahmedabad","SRH","KKR","SRH","KKR",160),
    ("Ahmedabad","RCB","RR","RR","RR",173),
    ("Chennai","SRH","RR","RR","SRH",176),
    ("Chennai","SRH","KKR","SRH","KKR",114)
]
```

Questions

- How many matches were played?
- What was the maximum run target?
- What was the average run target?
- How many matches had above average run targets?
- How many cities were venues?
- Which team played matches at maximum number of venues?
- Is winning the toss an advantage?

In [2]:
```
matchlist
```

Out[2]:
```
[('Chennai', 'RCB', 'CSK', 'RCB', 'CSK', 174),
 ('Mohali', 'DC', 'PK', 'PK', 'PK', 175),
 ('Kolkata', 'KKR', 'SRH', 'SRH', 'KKR', 209),
 ('Jaipur', 'RR', 'LSG', 'RR', 'RR', 194),
 ('Ahmedabad', 'GT', 'MI', 'MI', 'GT', 169),
 ('Bengaluru', 'PK', 'RCB', 'RCB', 'RCB', 177),
 ('Chennai', 'CSK', 'GT', 'GT', 'CSK', 207),
 ('Hyderabad', 'SRH', 'MI', 'MI', 'SRH', 278),
 ('Jaipur', 'RR', 'DC', 'DC', 'RR', 186),
 ('Bengaluru', 'RCB', 'KKR', 'KKR', 'KKR', 183),
 ('Lucknow', 'LSG', 'PK', 'LSG', 'LSG', 200),
 ('Ahmedabad', 'SRH', 'GT', 'SRH', 'GT', 163),
 ('Visakhapatnam', 'DC', 'CSK', 'DC', 'DC', 192),
 ('Mumbai', 'MI', 'RR', 'RR', 'RR', 126),
 ('Bengaluru', 'LSG', 'RCB', 'RCB', 'LSG', 182),
 ('Visakhapatnam', 'KKR', 'DC', 'KKR', 'KKR', 273),
 ('Ahmedabad', 'GT', 'PK', 'PK', 'PK', 200),
 ('Hyderabad', 'CSK', 'SRH', 'SRH', 'SRH', 166),
 ('Jaipur', 'RCB', 'RR', 'RR', 'RR', 184),
 ('Mumbai', 'MI', 'DC', 'DC', 'MI', 235),
 ('Lucknow', 'LSG', 'GT', 'LSG', 'LSG', 164),
 ('Chennai', 'KKR', 'CSK', 'CSK', 'CSK', 138),
 ('Mohali', 'SRH', 'PK', 'PK', 'SRH', 183),
 ('Jaipur', 'RR', 'GT', 'GT', 'GT', 197),
 ('Mumbai', 'RCB', 'MI', 'MI', 'MI', 197),
 ('Lucknow', 'LSG', 'DC', 'LSG', 'DC', 168),
 ('Mohali', 'PK', 'RR', 'RR', 'RR', 148),
 ('Kolkata', 'LSG', 'KKR', 'KKR', 'KKR', 162),
 ('Mumbai', 'CSK', 'MI', 'MI', 'CSK', 207),
 ('Bengaluru', 'SRH', 'RCB', 'RCB', 'SRH', 288),
 ('Kolkata', 'KKR', 'RR', 'RR', 'RR', 224),
 ('Ahmedabad', 'GT', 'DC', 'DC', 'DC', 90),
 ('Mohali', 'MI', 'PK', 'PK', 'MI', 193),
 ('Lucknow', 'CSK', 'LSG', 'LSG', 'LSG', 177),
 ('Delhi', 'SRH', 'DC', 'DC', 'SRH', 267),
 ('Kolkata', 'KKR', 'RCB', 'RCB', 'KKR', 223),
 ('Mohali', 'PK', 'GT', 'PK', 'GT', 143),
 ('Jaipur', 'MI', 'RR', 'MI', 'RR', 180),
 ('Chennai', 'CSK', 'LSG', 'LSG', 'LSG', 211),
 ('Delhi', 'DC', 'GT', 'GT', 'DC', 225),
 ('Hyderabad', 'RCB', 'SRH', 'RCB', 'RCB', 207),
 ('Kolkata', 'KKR', 'PK', 'PK', 'PK', 262),
 ('Delhi', 'DC', 'MI', 'MI', 'DC', 258),
 ('Lucknow', 'LSG', 'RR', 'RR', 'RR', 197),
 ('Ahmedabad', 'GT', 'RCB', 'RCB', 'RCB', 201),
 ('Chennai', 'CSK', 'SRH', 'SRH', 'CSK', 213),
 ('Kolkata', 'DC', 'KKR', 'DC', 'KKR', 154),
 ('Lucknow', 'MI', 'LSG', 'LSG', 'LSG', 145),
 ('Chennai', 'CSK', 'PK', 'PK', 'PK', 163),
 ('Hyderabad', 'SRH', 'RR', 'SRH', 'SRH', 202),
 ('Mumbai', 'KKR', 'MI', 'MI', 'KKR', 170),
 ('Bengaluru', 'GT', 'RCB', 'RCB', 'RCB', 148),
 ('Dharamsala', 'CSK', 'PK', 'PK', 'CSK', 168),
 ('Lucknow', 'KKR', 'LSG', 'LSG', 'KKR', 236),
 ('Mumbai', 'SRH', 'MI', 'MI', 'MI', 174),
 ('Delhi', 'DC', 'RR', 'RR', 'DC', 222),
 ('Hyderabad', 'LSG', 'SRH', 'LSG', 'SRH', 166),
 ('Dharamsala', 'RCB', 'PK', 'PK', 'RCB', 242),
 ('Ahmedabad', 'GT', 'CSK', 'CSK', 'GT', 232),
 ('Kolkata', 'KKR', 'MI', 'MI', 'KKR', 158),
 ('Chennai', 'RR', 'CSK', 'RR', 'CSK', 142),
 ('Bengaluru', 'RCB', 'DC', 'DC', 'RCB', 188),
 ('Delhi', 'DC', 'LSG', 'LSG', 'DC', 209),
 ('Guwahati', 'RR', 'PK', 'RR', 'PK', 145),
 ('Mumbai', 'LSG', 'MI', 'MI', 'LSG', 215),
 ('Bengaluru', 'RCB', 'CSK', 'CSK', 'RCB', 219),
```

```
('Hyderabad', 'PK', 'SRH', 'PK', 'SRH', 215),
('Ahmedabad', 'SRH', 'KKR', 'SRH', 'KKR', 160),
('Ahmedabad', 'RCB', 'RR', 'RR', 'RR', 173),
('Chennai', 'SRH', 'RR', 'RR', 'SRH', 176),
('Chennai', 'SRH', 'KKR', 'SRH', 'KKR', 114)]
```

How many matches were played?

- Python's built-in function `len()` directly tells the length of a list

In [3]:
```
len(matchlist)
```

Out[3]: 71

Count the number of entries explicitly

- Iterate through the list
- `for` variable `in` listname `:`
  - `:` indicates a block of statements (commands) to follow
  - Indent the commands to be performed within each iteration
- `count = count + 1`
  - `=` assigns a new value to a variable, not to be confused with equality
  - rhs is old value of `count`, used to update the value of `count`

In [4]:
```
count = 0
for row in matchlist:
    count = count + 1  # Compute current count + 1 and reassign it to count
```

In [5]:
```
count
```

Out[5]: 71

- Python does not require you to "declare" variables in advance
- Can introduce new names on the fly
- This can be a source of errors, if you mistype a name

In [6]:
```
count = 0
for row in matchlist:
    countt = count + 1  # Compute current count + 1 and reassign it to count
```

In [7]:
```
count
```

Out[7]: 0

- If you are lucky, the mistyped name will appear on the right and Python will flag an error

In [8]:
```
count = 0
for row in matchlist:
    count = ccount + 1  # Compute current count + 1 and reassign it to count
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
Cell In[8], line 3
      1 count = 0
      2 for row in matchlist:
----> 3     count = ccount + 1  # Compute current count + 1 and reassign it to count

NameError: name 'ccount' is not defined
```

What was the maximum run target?

- Initialize `maxtarget` to `-1`
- Update `maxtarget` whenever run target in current row exceeds current maximum
  - Extract run target from tuple of values of current row using positional index
  - In Python, indices always start with 0, so six entries in the tuple have indices 0,1,2,3,4,5
- `if` specifies conditional execution
  - `if` conditional-expression `:`
    - Expression evaluates to `True` or `False`
    - As before, `:` and indentation indicate scope of conditional execution
    - Indented commands (update of `maxtarget`) happens only when `if` condition evaluates to `True`

In [9]:
```
maxtarget = -1
for row in matchlist:
    target = row[5] # Sixth component, indexed 0,1,..,5
    if target > maxtarget:
        maxtarget = target
```

In [10]:
```
maxtarget
```

Out[10]:
```
288
```

- No need to assign a separate variable for `row[5]

In [11]:
```
maxtarget = -1
for row in matchlist:
    if row[5] > maxtarget:
        maxtarget = row[5]
```

In [12]:
```
maxtarget
```

Out[12]:
```
288
```

- Can also initialize `maxtarget` to first target in the table
- Not a problem to read the first row of the table again

In [13]:
```
firstrow = matchlist[0]
maxtarget = firstrow[5]
for row in matchlist:  # Not a problem to read first row again
    if row[5] > maxtarget:
        maxtarget = row[5]
```

What was the average run target?

- Iterate to compute sum of all run targets

- Average is total divided by count

In [14]:
```
count = 0
targetsum = 0
for row in matchlist:
    count = count+1
    targetsum = targetsum + row[5]
targetavg = targetsum/count
```

In [15]:
```
targetavg
```

Out[15]: 190.59154929577466

- Can also maintain running average instead of running total

In [16]:
```
count = 0
currentavg = 0
for row in matchlist:
    totalsofar = count*currentavg
    count = count+1
    currentavg = (totalsofar + row[5])/count
```

In [17]:
```
currentavg
```

Out[17]: 190.59154929577466

- Alternatively, avoid creating running total altogether

In [18]:
```
count = 0
currentavg = 0
for row in matchlist:
    count = count+1
    currentavg = (currentavg/count)*(count-1) + row[5]/count
```

In [19]:
```
currentavg
```

Out[19]: 190.59154929577463

How many matches had above average run targets?

- *Filtered* iteration
- Update the count only if the current run target is $\geq$ average

In [20]:
```
count = 0
targetsum = 0
for row in matchlist:
    count = count+1
    targetsum = targetsum + row[5]
targetavg = targetsum/count

aboveavgcount = 0
for row in matchlist:
    if row[5] > targetavg:
        aboveavgcount = aboveavgcount + 1
```

In [21]:
```
aboveavgcount
```

Out[21]: 34

How many cities were venues?

- Maintain a list of venues
  - Need to initialize to empty list to tell Python this value can be used as a list
  - Built-in check `v in l` returns `True` if value `v` is present in list `l`
  - Add new venue to the list -- `l1 + l2` concatenates two lists into a single list

In [22]:
```
venuelist = []
for row in matchlist:
    venue = row[0]  # Leftmost column
    if not(venue in venuelist):
        venuelist = venuelist + [venue]
```

In [23]:
```
venuelist
```

Out[23]:
```
['Chennai',
 'Mohali',
 'Kolkata',
 'Jaipur',
 'Ahmedabad',
 'Bengaluru',
 'Hyderabad',
 'Lucknow',
 'Visakhapatnam',
 'Mumbai',
 'Delhi',
 'Dharamsala',
 'Guwahati']
```

In [24]:
```
len(venuelist)
```

Out[24]:
```
13
```

- Can instead use a *dictionary*
  - A collection of key:value pairs
  - Initialize to empty dictionary
  - Check if venue already exists as a key
    - Create a new key if the current venue is not a key
    - Value assigned to the key is unimportant here

In [25]:
```
venuedict = {}
for row in matchlist:
    venue = row[0]  # Leftmost column
    if not(venue in venuedict.keys()):
        venuedict[venue] = "Something"
```

In [26]:
```
venuedict
```

Out[26]:
```
{'Chennai': 'Something',
 'Mohali': 'Something',
 'Kolkata': 'Something',
 'Jaipur': 'Something',
 'Ahmedabad': 'Something',
 'Bengaluru': 'Something',
 'Hyderabad': 'Something',
 'Lucknow': 'Something',
 'Visakhapatnam': 'Something',
 'Mumbai': 'Something',
 'Delhi': 'Something',
 'Dharamsala': 'Something',
 'Guwahati': 'Something'}
```

- Can extract keys of a dictionary
  - Similar to, but not quite a list

In [27]:
```
venuedict.keys()
```

Out[27]:
```
dict_keys(['Chennai', 'Mohali', 'Kolkata', 'Jaipur', 'Ahmedabad', 'Bengaluru', 'Hy
derabad', 'Lucknow', 'Visakhapatnam', 'Mumbai', 'Delhi', 'Dharamsala', 'Guwahat
i'])
```

- By convention, Python assumes you mean `d.keys()` if you say `v in d` for a dictionary `d`

In [28]:
```
venuedict2 = {}
for row in matchlist:
    venue = row[0]  # Leftmost column
    if not(venue in venuedict2):  # Short form for venue in venuedict2.keys()
        venuedict2[venue] = "Something"
```

In [29]:
```
venuedict2
```

Out[29]:
```
{'Chennai': 'Something',
 'Mohali': 'Something',
 'Kolkata': 'Something',
 'Jaipur': 'Something',
 'Ahmedabad': 'Something',
 'Bengaluru': 'Something',
 'Hyderabad': 'Something',
 'Lucknow': 'Something',
 'Visakhapatnam': 'Something',
 'Mumbai': 'Something',
 'Delhi': 'Something',
 'Dharamsala': 'Something',
 'Guwahati': 'Something'}
```

- Can also count the number of matches at each venue
- For each key, maintain a counter
  - When the key is created, initialize the counter to 1
  - If the key already exists, increment the counter

In [30]:
```
venuedict3 = {}
for row in matchlist:
    venue = row[0]  # Leftmost column
    if not(venue in venuedict3):  # Short form for venue in venuedict3.keys()
        venuedict3[venue] = 1
    else:
        venuedict3[venue] = venuedict3[venue] + 1
```

In [31]:
```
venuedict3
```

Out[31]:
```
{'Chennai': 9,
 'Mohali': 5,
 'Kolkata': 7,
 'Jaipur': 5,
 'Ahmedabad': 8,
 'Bengaluru': 7,
 'Hyderabad': 6,
 'Lucknow': 7,
 'Visakhapatnam': 2,
 'Mumbai': 7,
 'Delhi': 5,
 'Dharamsala': 2,
 'Guwahati': 1}
```

- Can invert the condition and exchange the two parts of the `if`

In [32]:
```
venuedict4 = {}
for row in matchlist:
    venue = row[0]  # Leftmost column
    if venue in venuedict4:  # Short form for venue in venuedict4.keys()
        venuedict4[venue] = venuedict4[venue] + 1
    else:
        venuedict4[venue] = 1
```

In [33]:
```
venuedict4
```

Out[33]:
```
{'Chennai': 9,
 'Mohali': 5,
 'Kolkata': 7,
 'Jaipur': 5,
 'Ahmedabad': 8,
 'Bengaluru': 7,
 'Hyderabad': 6,
 'Lucknow': 7,
 'Visakhapatnam': 2,
 'Mumbai': 7,
 'Delhi': 5,
 'Dharamsala': 2,
 'Guwahati': 1}
```

Which team played matches at maximum number of venues?

- Maintain a dictionary where keys are teams
- For each team, associate a list of venues where it has played
- Each match involves two teams, so update the dictionary for both teams

In [34]:
```
teamvenuedict = {}
for row in matchlist:
    venue = row[0]
    team1 = row[1]
    team2 = row[2]

    # Update the dictionary for team1
    if not (team1 in teamvenuedict):
        teamvenuedict[team1] = [venue]
    else:
        if not (venue in teamvenuedict[team1]):
            teamvenuedict[team1] = teamvenuedict[team1] + [venue]
```

```
    # Update the dictionary for team2
    if not (team2 in teamvenuedict):
        teamvenuedict[team2] = [venue]
    else:
        if not (venue in teamvenuedict[team2]):
            teamvenuedict[team2] = teamvenuedict[team2] + [venue]
```

In [35]:
```
teamvenuedict
```

Out[35]:

```
{'RCB': ['Chennai',
  'Bengaluru',
  'Jaipur',
  'Mumbai',
  'Kolkata',
  'Hyderabad',
  'Ahmedabad',
  'Dharamsala'],
 'CSK': ['Chennai',
  'Visakhapatnam',
  'Hyderabad',
  'Mumbai',
  'Lucknow',
  'Dharamsala',
  'Ahmedabad',
  'Bengaluru'],
 'DC': ['Mohali',
  'Jaipur',
  'Visakhapatnam',
  'Mumbai',
  'Lucknow',
  'Ahmedabad',
  'Delhi',
  'Kolkata',
  'Bengaluru'],
 'PK': ['Mohali',
  'Bengaluru',
  'Lucknow',
  'Ahmedabad',
  'Kolkata',
  'Chennai',
  'Dharamsala',
  'Guwahati',
  'Hyderabad'],
 'KKR': ['Kolkata',
  'Bengaluru',
  'Visakhapatnam',
  'Chennai',
  'Mumbai',
  'Lucknow',
  'Ahmedabad'],
 'SRH': ['Kolkata',
  'Hyderabad',
  'Ahmedabad',
  'Mohali',
  'Bengaluru',
  'Delhi',
  'Chennai',
  'Mumbai'],
 'RR': ['Jaipur',
  'Mumbai',
  'Mohali',
  'Kolkata',
  'Lucknow',
  'Hyderabad',
  'Delhi',
  'Chennai',
  'Guwahati',
  'Ahmedabad'],
 'LSG': ['Jaipur',
  'Lucknow',
  'Bengaluru',
  'Kolkata',
  'Chennai',
  'Hyderabad',
  'Delhi',
```

```
 'Mumbai'],
'GT': ['Ahmedabad',
 'Chennai',
 'Lucknow',
 'Jaipur',
 'Mohali',
 'Delhi',
 'Bengaluru'],
'MI': ['Ahmedabad',
 'Hyderabad',
 'Mumbai',
 'Mohali',
 'Jaipur',
 'Delhi',
 'Lucknow',
 'Kolkata']}
```

Is winning the toss an advantage?

- Count the rows where the toss winner is the match winner
- Check if the fraction of such rows is above a given threshold

In [36]:
```
threshold = 2/3
count = 0
tossandwin = 0
for row in matchlist:
    count = count + 1
    tosswinner = row[3]
    matchwinner = row[4]
    if tosswinner == matchwinner: # Use == for equality check
        tossandwin = tossandwin + 1
```

- Is the ratio above the threshold?

In [37]:
```
tossandwin/count >= threshold
```

Out[37]:
```
False
```

In [38]:
```
tossandwin/count
```

Out[38]:
```
0.43661971830985913
```