

Data warehouse and Data Mining (TDT 4300)

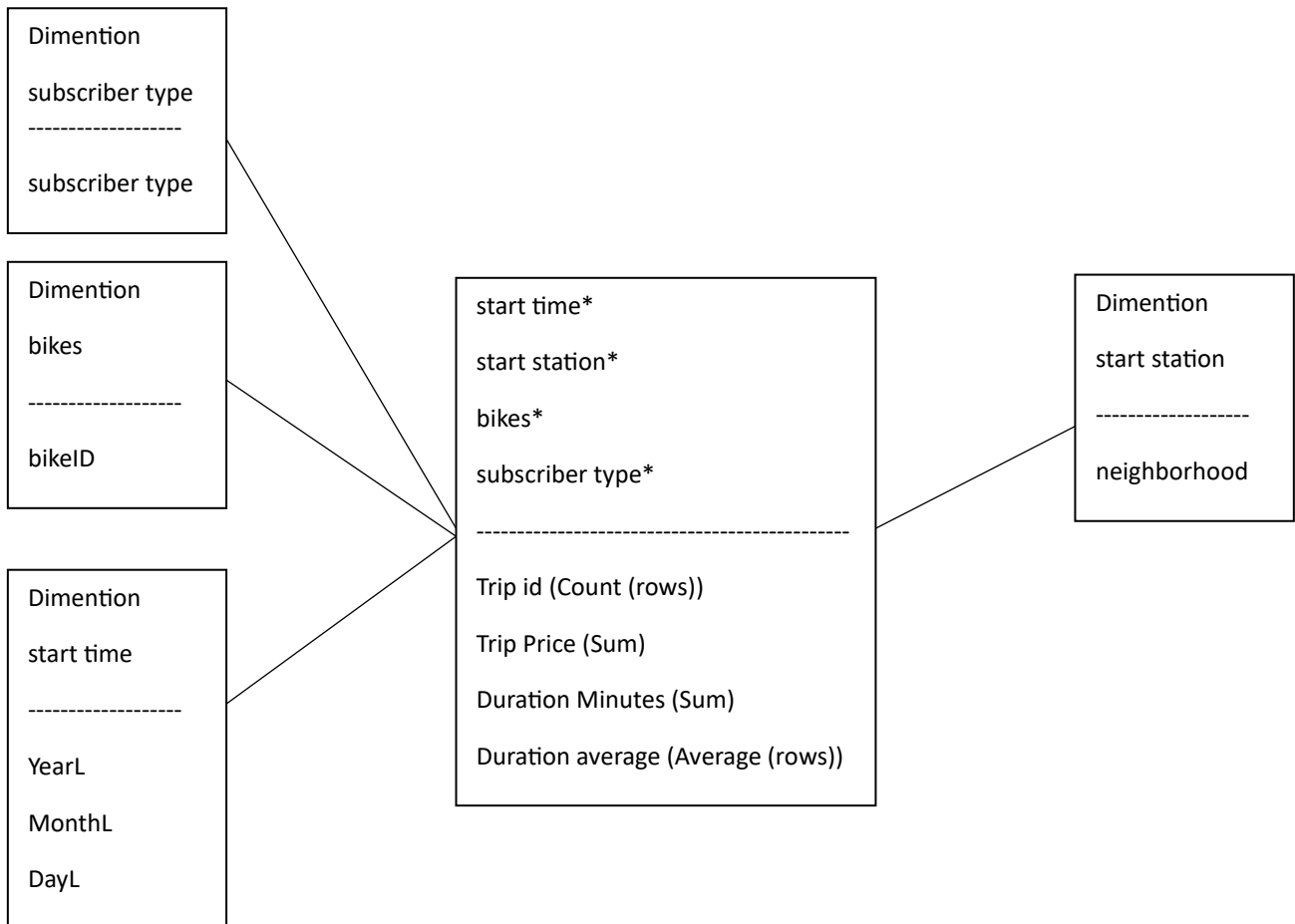
Assignment 1

Submitted by

Name : Md Anwarul Hasan
Student ID : 583233

Part 1

The schema:



The Assumption:

From the reports, I have realized that I will require below mentioned dimensions,

Serial number	Dimensions	Reason
1	start time with level YearL, MonthL and DayL	Because every report needed to work with years and in addition to that Report 3 required months as well. I brought it down to even days as I thought it may be required although I later realized I did not have break it down to this granularity.
2	bikes with level bikeID	Report 2 needed that.
3	subscriber type with level subscriber type	Report 4 needed that.
4	start station with level neighborhood	Report 5 needed that

Report 1: Trip counts in each year.

MDX Query:

```
select [Measures].[Trip Id] on columns,  
      [start time].[start time].[YearL] on rows  
from [bikeTCube]
```

Snapshot:

	Trip Id
➤ 2014	13,273.00
➤ 2016	18,561.00
➤ 2017	8,963.00
➤ 2015	17,365.00
➤ 2013	139.00

Report 2: The total duration each bike was ridden in 2014.

MDX Query:

```
select  
[Measures].[Duration Minutes] on columns,  
[bikes].[bikes].[bikeID]on rows  
from [bikeTCube]  
where [start time].[start time].[All].&[2014]
```

Snapshot:

(the result consists lot of rows. For that reason, I exported the result in excel and from the excel copied the result and pasted here.)

	Duration Minutes
58	1141
696	1322
203	613
157	1014
894	660
473	650

	Duration Minutes
52	1766
25	1105
411	1507
528	1149
352	1939
440	337
135	1237
209	1923
478	444
775	
559	718
276	2853
545	817
976	1382
429	1138
97	639
503	1629
36	1218
226	1144
248	681
924	961
747	220
979	493
870	
298	1203
26	1256
484	1141
14	666
669	706
488	620
445	926
393	381
275	1300
983	2298
301	469
285	875
371	1361
64	946
235	1015
961	837
344	863
321	786
800	1105
541	488

	Duration Minutes
882	417
934	1171
387	1083
233	
646	1417
191	2845
685	1472
320	900
35	1491
451	
334	966
709	
856	306
897	724
20	1091
219	545
579	301
398	943
208	1436
102	1318
221	900
241	855
843	1215
850	1150
852	1126
160	625
563	205
867	153
140	1152
406	1012
974	1160
182	1696
930	437
105	553
84	1070
113	1091
303	1097
513	896
902	930
553	1081
853	355
514	
299	2037
683	721

	Duration Minutes
981	2045
751	196
517	806
152	887
247	557
287	1667
75	954
354	569
190	1961
277	710
425	933
282	1125
362	567
479	895
142	1104
223	1164
963	845
88	1111
442	1156
861	223
104	628
446	1363
158	927
230	1585
801	323
106	718
561	
213	884
789	
874	
4	
969	1198
100	864
428	1262
466	1037
270	506
395	2215
197	1199
40	1902
854	1547
341	538
399	1729
326	1389
332	1624

	Duration Minutes
434	732
712	2725
588	1295
881	332
410	1682
60	1073
93	889
288	1604
206	708
610	176
839	1627
212	409
650	2189
991	796
716	814
805	934
22	627
382	250
827	1666
975	1674
970	1782
885	680
27	816
15	1085
995	192
28	520
956	1041
207	2476
116	1187
373	563
472	
417	1355
195	546
896	1549
266	801
150	1761
936	769
421	961
193	896
420	621
5089	
188	753
29	1167
8	748

	Duration Minutes
666	
311	1061
80	1237
957	736
555	4257
453	907
340	764
114	1438
690	1063
439	789
823	1054
272	1396
55	740
746	1624
164	943
369	474
200	679
124	1623
41	1113
169	922
815	121
107	1179
132	710
327	1100
668	1246
254	1346
951	1339
68	1120
214	1844
432	1423
72	1400
822	376
971	416
460	357
384	682
911	732
520	487
871	
284	669
161	1115
177	1688
664	537
765	230
577	

	Duration Minutes
240	736
409	1121
572	752
984	1244
356	1510
391	264
110	2046
621	1245
980	1217
122	593
958	1395
745	209
575	1526
661	1054
370	928
865	363
855	
890	433
19	800
262	1050
263	1566
860	172
928	254
79	1724
589	
968	2108
386	1482
749	328
547	564
548	1055
198	1490
133	1089
947	1106
468	820
265	
625	713
546	
922	1920
893	391
578	
938	1804
772	892
407	3468
884	168

	Duration Minutes
95	3709
937	715
85	1005
24	406
127	598
869	1398
551	1342
670	191
933	739
925	1079
965	577
920	1609
153	595
593	1716
999	1459
878	934
863	128
910	1460
558	944
252	484
16	911
3	2274
101	1239
134	1338
416	739
367	1124
647	1205
464	441
54	1700
465	313
540	1542
706	906
2459	
86	988
655	
181	769
281	1821
554	1190
224	1670
397	723
955	522
898	846
539	952
166	

	Duration Minutes
360	986
996	203
183	835
735	543
37	840
459	696
660	1516
907	1257
78	330
228	1256
400	435
229	1699
46	664
141	843
866	152
522	876
895	1757
414	1603
49	1628
396	871
509	689
178	1629
165	458
864	233
1000	1609
518	
550	1290
461	986
680	614
273	757
571	
998	760
348	1615
849	48
873	
529	713
470	455
919	511
759	1028
202	919
511	1939
877	547
175	1031
962	3038

	Duration Minutes
347	917
185	960
62	528
329	1228
523	1113
302	2424
306	906
176	1046
315	362
966	909
236	1051
232	1872
576	1230
872	
89	1486
544	1184
5	
875	
201	1461
829	941
349	1017
447	1344
774	1310
30	1473
283	1278
385	183
549	900
412	1036
260	1346
117	1089
560	682
435	1873
2147	
507	197
693	
452	360
220	1263
422	728
242	381
728	
663	379
388	
851	120
483	1504

	Duration Minutes
993	769
862	143
832	
2048	
125	969
172	30
988	334
987	3

Report 3: The trip count, revenue and average trip duration throughout the months from 2013 to 2015. Can you observe some patterns?

MDX Query:

select

{[Measures].[Trip Id], [Measures].[Trip Price], [Measures].[Duration average]} on 0,
 {[start time].[start time].[All].&[2013],
 [start time].[start time].[All].&[2014],
 [start time].[start time].[All].&[2015]} on rows
 from [bikeTCube]

Snapshot:

	Trip Id	Trip Price	Duration avera...
▼ 2013	139.00	513.05	47.15827338129
▶ 12	139.00	513.05	47.15827338129
▼ 2014	13,273.00	30,234.63	29.10381978452
▶ 3	2,284.00	5,696.38	31.86471103327
▶ 12	647.00	1,189.55	23.49304482225
▶ 5	1,292.00	3,072.47	30.38312693498
▶ 4	1,263.00	3,099.87	31.35708630245
▶ 10	1,628.00	3,471.37	27.24324324324
▶ 8	937.00	2,026.86	27.63820704375
▶ 7	1,126.00	2,421.71	27.48046181172
▶ 2	530.00	1,367.08	32.95660377358
▶ 11	1,010.00	2,063.91	26.10792079207
▶ 6	1,203.00	3,200.96	33.99501246882
▶ 1	328.00	746.57	29.08231707317
▶ 9	1,025.00	1,877.90	23.40878048780
▼ 2015	17,365.00	35,855.95	26.38186006334

➤ 6	1,301.00	2,573.17	25.26979246733
➤ 5	1,510.00	3,486.28	29.49801324503
➤ 2	798.00	1,663.77	26.63784461152
➤ 9	1,536.00	2,878.65	23.94466145833
➤ 8	1,289.00	2,147.51	21.28704422032
➤ 12	981.00	2,862.41	37.28032619775
➤ 10	2,389.00	5,031.59	26.90916701548
➤ 4	1,531.00	3,437.37	28.68582625734
➤ 3	2,622.00	4,828.34	23.52898550724
➤ 7	1,487.00	2,895.39	24.87895090786
➤ 11	1,214.00	2,635.77	27.73970345963
➤ 1	707.00	1,415.70	25.58415841584

Pattern:

Both the year 2014 & 2015 have the month 3 as with highest total trip and with highest accumulated trip price. And month like 4 to 11 almost always made total price more than 1000 on 2014 & 2015 as well. Average trip duration is somewhat near 25 almost all the time.

Report 4: Subscriber types that generated revenue greater than 400 (USD) in the year 2015.

MDX Query:

select

[Measures].[Trip Price] on columns,

filter([subscriber type].[subscriber type].[subscriber type], [Measures].[Trip Price] > 400)
on rows

from [bikeTCube]

where [start time].[start time].[All].&[2015]

Snapshot:

	Trip Price
Walk Up	29,072.41
Local30	1,169.09
Local365	3,783.06
Weekender	622.19
Explorer	596.49

Report 5: The share of trips per neighborhood and per year. Consider only years 2014 and 2015.

MDX Query:

with

member

[Trip share] as [Measures].[Trip Id] / sum(axis(1), [Measures].[Trip Id]),
FORMAT_STRING = 'percent'

select

([Trip share], [start time].[start time].[YearL].members) on columns,

[start station].[start station].[neighborhood] on rows

from (SELECT {[start time].[start time].[All].&[2014]}, [start time].[start time].[All].&[2015]} ON COLUMNS FROM [bikeTCube])

Snapshot:

	Trip share	
	➤ 2014	➤ 2015
➤ Downtown	62.81%	60.67%
➤ Barton Hills	2.14%	3.87%
➤ University of...	0.00%	0.00%
➤ Bouldin	14.56%	11.36%
➤ East Cesar ...	4.65%	4.60%
➤ West Univer...	3.65%	2.98%
➤ Old West Au...	0.00%	0.09%
➤ South River ...	2.81%	2.54%
➤ Zilker	7.00%	8.79%
➤ Holly	0.74%	2.29%
➤ Central East ...	1.45%	2.38%
➤ Windsor Ro...	0.00%	0.00%
➤ Govalle	0.20%	0.43%