

# INFSCI 2710 – Database Management – Spring 2018

## Homework 5 – A Data-Driven Web Application

1. Use the settings in lab session to connect to your MySQL and phpmyadmin
2. For all tasks, use the **Yelp database (same as Homework #4)** to fulfill the requirements.
3. Name the Word document (or PDF as I recommended) with your answers as YourPittID\_infsci2710\_homework5.docx (pdf). In other words, if your Pitt ID (first part of your Pitt email) is abc123, your submission file should be named abc123\_infsci2710\_homework5.docx (pdf)
4. Zip both your code folder and document as YourPittID\_infsci2710\_homework5.zip
5. Submit the Zip file via CourseWeb

**Task1 (25 points):** Create two in your yelp\_db database. The procedure should

- 1) reproduce the answer of “Task 6 in your Homework #4”;
- 2) Add a usage record to a newly created table – Social\_Function\_Usage (see the table below). The table will be inserted a new row ([id], [procedure name], [date]) while calling the stored procedure.

**procedure 1:** Find top 10 restaurants which got the most reviews in a given date.

```
CREATE DEFINER=`root`@`localhost` PROCEDURE `daily_topreview_rest` (IN `date_time` DATETIME) NOT DETERMINISTIC READS SQL DATA SQL SECURITYDEFINER
BEGIN
SELECT b.id business_id, b.name business_name FROM business b JOIN (SELECT id, business_id
FROM review WHERE date = date_time) r ON b.id=r.business_id JOIN(SELECT*FROM category WHERE category='Restaurants') c ON b.id=c.business_id GROUP BY b.id ORDER BY COUNT(r.id) DESC LIMIT 10;
INSERT INTO social_function_usage (call_function,date) VALUES ('daily_topreview_rest', CURRENT_DATE);
END
```

✓ Routine 'daily\_topreview\_rest' has been modified.

DROP PROCEDURE `daily\_topreview\_rest` ; CREATE DEFINER='root'@'localhost' PROCEDURE `daily\_topreview\_rest` (IN `date\_time` DATETIME) NOT DETERMINISTIC READS SQL DATA SQL SECURITY DEFINER BEGIN SELECT b.id business\_id, b.name business\_name FROM business b JOIN (SELECT id, business\_id FROM review WHERE date = date\_time) r ON b.id=r.business\_id JOIN (SELECT\*FROM category WHERE category='Restaurants') c ON b.id=c.business\_id GROUP BY b.id ORDER BY COUNT(r.id) DESC LIMIT 10; INSERT INTO social\_function\_usage (call\_function, date) VALUES ('daily\_topreview\_rest', CURRENT\_DATE); END

[Edit inline] [Edit] [Create PHP code]

Routines
Name Action Type Returns key_word_business Edit Execute Export Drop PROCEDURE

**procedure 2:** Find Coffee & Tea shop with free WiFi in a given city.

```
CREATE PROCEDURE `city_wifi_coffee` (IN `city_name` VARCHAR(255) CHARSET big5) NOT DETERMINISTIC READS SQL DATA SQL SECURITY DEFINER
BEGIN
SELECT DISTINCT b.id, b.name FROM(select id, name from business where city=city_name) b JOIN (select business_id from attribute where name='WiFi' AND value='free') a ON b.id=a.business_id JOIN (select business_id from category where category='Coffee & Tea') c ON b.id=c.business_id;
INSERT INTO social_function_usage (call_function, date) VALUES ('city_wifi_coffee',CURRENT_DATE);
END
```

✓ Routine 'city\_wifi\_coffee' has been created.

```
CREATE PROCEDURE `city_wifi_coffee`(IN `city_name` VARCHAR(255) CHARSET big5) NOT DETERMINISTIC READS SQL DATA SQL SECURITY DEFINER BEGIN SELECT DISTINCT b.id, b.name FROM (select id, name from business where city=city_name) b JOIN (select business_id from attribute where name='WiFi' AND value='free') a ON b.id=a.business_id JOIN (select business_id from category where category='Coffee & Tea') c ON b.id=c.business_id; INSERT INTO social_function_usage (call_function, date) VALUES ('city_wifi_coffee', CURRENT_DATE); END
```

[Edit inline] [ Edit ] [ Create PHP code ]

Routines					
Name	Action	Type	Returns		
<input type="checkbox"/> daily_topreview_rest		PROCEDURE			
<input type="checkbox"/> key_word_business		PROCEDURE			
<input type="checkbox"/> city_wifi_coffee		PROCEDURE			
<input type="checkbox"/> Check all	<input type="checkbox"/> With selected:				

```
CREATE TABLE social_function_usage (
    id INT PRIMARY KEY NOT NULL AUTO_INCREMENT,
    call_function VARCHAR(20) NOT NULL,
    date DATETIME NOT NULL
)
```

✓ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0211 seconds.)

```
CREATE TABLE social_function_usage( id INT PRIMARY KEY NOT NULL AUTO_INCREMENT, call_function VARCHAR(20) NOT NULL, date DATETIME NOT NULL )
```

[Edit inline] [ Edit ] [ Create PHP code ]

Field name	Primary Key	Data type	Is null?	Is auto-increment?
Id	yes	INT	No	yes
Call Function	No	VARCHAR(20)	No	no
Date	No	Datetime	No	no

## Call procedures:

Execution results of routine `city_wifi_coffee`	
Your SQL query has been executed successfully. 10 rows affected by the last statement inside the procedure.	
SET <u>@p0='2016-08-07'</u> ; <u>CALL `daily_topreview_rest`(@p0);</u>	
Execution results of routine `daily_topreview_rest`	
Your SQL query has been executed successfully. 10 rows affected by the last statement inside the procedure.	
SET <u>@p0='2016-08-07'</u> ; <u>CALL `city_wifi_coffee`(@p0);</u>	
Execution results of routine `city_wifi_coffee`	
Your SQL query has been executed successfully. 198 rows affected by the last statement inside the procedure.	
SET <u>@p0='Pittsburgh'</u> ; <u>CALL `city_wifi_coffee`(@p0);</u>	

✓ Your SQL query has been executed successfully.  
198 rows affected by the last statement inside the procedure.

```
SET @p0='Pittsburgh'; CALL `city_wifi_coffee`(@p0);
```

id	name
-8F04F54IDT6VgWPCgybug	Ineffable Ca Phe
-CfqFIH6rxUxJL9gR-k_XA	Dunkin' Donuts
-SUammimmS2p6c5GuuGA-kmw	Buon Giorno Cafe
0bjFYstj8USMzEV4ZQldja	Redhawk Coffee
0lpT5qj09vUfDp5zUBwjlw	Commonplace Coffee
0MR1J0l8uOxIvhZIHv-johQ	Kickback Pinball Cafe
0PJQ-WR-S286xH6m2RWWeQw	Starbucks
1eJ7ylz5AjwC145nHD5Xw	Starbucks
1h1EGiDCoKQnL-VmRHlagw	Starbucks
2-FFIBaAN2nZhck-K-GBR75w	Crazy Mocha
2-ijqX-PonxJUSZGugyCOAQ	Crazy Mocha
2cF-sCqhgCGzVMJULhNqjBQ	Dunkin' Donuts
2M8HnQRsKqETrmx8YhtvGQ	Commonplace Voluto
3GuqqEYaf1dWXcwOX5xPw	Square Cafe
48nQpsNIVcnTR0nXxs9gyA	Trolley Barn Eatery
4AgSOPmhRPcC-NctjszQ	Peet's Coffee & Tea
4LileRb...r9pUCM0mKS9XA	Friendship Perk And Brew

**Task 2 (25 points):** Write a PHP application to call the two stored procedures. You need to itemize all return results. The user should be able to click and inspect the details. For example, in the sample social function of 'top restaurant of the day', a user can 1) browse the restaurants by different given date;

2) click the link in the browsing page to inspect the detail of the restaurant.

**Procedure1:**

## inputInfo1.php

The screenshot shows a Chrome browser window with the URL `localhost:8888/jjc117_infsci2710_homework5/inputInfo.php`. The page displays a message: "Connected successfully" followed by "Find top 10 restaurants which got the most reviews in a given date." Below this, there is a text input field labeled "Please input a date in a format of 'yyyy-mm-dd'." with the value "Date: 2016-08-07". A "submit" button is located below the input field. The browser's toolbar and a vertical toolbar on the left side of the screen are visible.

## browsingPage1.php

The screenshot shows a Chrome browser window with the URL `localhost:8888/jjc117_infsci2710_homework5/browsingPage1.php`. The page displays a message: "Connected successfully" followed by a table with two columns: "business\_id" and "business\_name". The table contains 10 rows of data. The browser's toolbar and a vertical toolbar on the left side of the screen are visible.

business_id	business_name
cYwJ1A2A6l12KNkm2rtXd5g	Gordon Ramsay BurGR
OMKgn9p_hRUiumBDAAFWIA	Cutting Board Filipino Gastropub
uh7DUWtpPoZkuEE05fghJ_w	Beauty & Essex
0qpNcqluKE88rKhbSZDpg	Killer Shrimp
YVd2yYNjnpifg2s4ToMRmPA	Cinnaholic
ysv6yhYYOoI9P7PlMyD0g	Wahlburgers
JDZ6_yycNQfTpUzJLKHUp	Ei Dorado Cantina
XQI9r_jZR1vOxZVcZfDDwQ	Cracker Barrel Old Country Store
k_tg18rWRibz4jKpqqhB0Q	Bonchon Chicken
OtSbxxD-WNzlNQjDGpXESA	IKEA

## detailPage.php (procedure1)

A screenshot of a Mac OS X desktop with a Chrome browser window open. The URL is `localhost:8888/jic117_infsci2710_homework5/detailPage.php?businessID=OtSbxxD-WNzINQIDGpXFSA`. The browser status bar shows "Connected successfully". The main content area displays a table with one row of data:

id	name	neighborhood	address	city	state	postal_code	latitude	longitude	stars	review_count	is_open
OtSbxxD-WNzINQIDGpXFSA	IKEA	Spring Valley	6500 Ikea Way	Las Vegas	NV	89148	36.0689	-115.281	3.5	303	1

If the system do not find the relevant data, the page will show like this screen shot.

A screenshot of a Mac OS X desktop with a Chrome browser window open. The URL is `localhost:8888/jic117_infsci2710_homework5/browsingPage1.php`. The browser status bar shows "Connected successfully". The main content area displays the message: "No results have been returned. [Please try another date.](#)"

## Procedure2: inputInfo2.php

Find Coffee & Tea shop with free WiFi in a given city.

Please input a city name.

City:

## browsingPage2.php

business_id	business_name
-8F04FS4dDT6VgWPcgybug	Ineffable Ca Phe
-CfqFH6rxUxI9gR-k_XA	Dunkin' Donuts
-SUammnS2p6c5GuuGA-kmw	Buon Giorno Cafe
0bJFYsJ8USMzEV4ZQldjA	Redhawk Coffee
0jpT5sJ09yuJDP5zLlBwjLw	Commonplace Coffee
0MR1J0U8uOXtvZIHv-johQ	Kickback Pinball Cafe
0PJ0-WR-S286xH6m2RWeQw	Starbucks
1e7ylsAjsxC145fnHD5Xw	Starbucks
1h1EGidCoKQnL-VmRHlagw	Starbucks
2-FFFbaaN2n7eK-GBR75w	Crazy Mocha
2-jyaX-PonxUSZQugyCOAQ	Crazy Mocha
2cF-cqhCGzVMJULhNqjBQ	Dunkin' Donuts
2M8HbQRsKqETmxo8YhTvGQ	Commonplace Volute
3GuqcEYaF1d]WXcwOX5xPw	Square Cafe
48nPqsNiVen1TRUnXsg7yA	Trolley Barn Eatery
4AgSOPmhRPc-NctjszQ	Peet's Coffee & Tea
4LlieRb- r9pUCM0mKS9XA	Friendship Perk And Brew
4mb7Wj8RAzCxSCRt4VSjhA	Perry Perk
SerOr-2phHfFeGOrfeROQA	Brunoise Restaurant
5h7ra0kJKswSZezbFSIpan	Beehive Coffee
5kA4hmzNNWGGrwvoGnlyr	Tango Cafe
5u27mhDjc4Y7QbDhRyaXg	Fuku Tea
61NPgdR8sMwAWwOC178nBg	Starbucks
6bR-79luBagZlbgLFWg	Amani Coffee House & Cafe
6TlqrH6HWpkitX-O2VmZXw	T.A.U.
7945an0vXONmkZ1n7Mu7Q	21st Street Coffee And Tea
7Wt1j4ppNwGyrN6PChpqQ	Dunkin' Donuts
87rlshpXfYcJXN1_DiUQ9A	Crazy Mocha
8vqAI5XLU8cQc38pbhE0Xw	Starbucks
005-cDm3uFu4mu4N of Knw	Starbucks

## detailPage.php (same php file as procedure1)

id	name	neighborhood	address	city	state	postal_code	latitude	longitude	stars	review_count	is_open
ObjFYstj8USMzEV4ZQldjA	Redhawk Coffee	Oakland	120 Meyran Ave	Pittsburgh	PA	15213	40.4407	-79.9586	5	91	1

**Task 3 (25 points):** Write a search function (search.php) so that the user can search (filter) the businesses through the review text in the Yelp data. The user should be able to search the businesses with a specific keyword, e.g., to explore the businesses with the keyword ‘awesome’ in their reviews. You need to itemize all return results. The user should be able to click and inspect the details.

```
CREATE DEFINER=`root`@`localhost` PROCEDURE `key_word_business`(`IN`  
`keyword` VARCHAR(10)) NOT DETERMINISTIC READS SQL DATASQL SECURITY DEFINER SELECT DISTINCT b.id, b.name FROM business b, (select business_id, text from review where MATCH (text)  
AGAINST (keyword)) r WHERE b.id =r.business_id
```

In order to improve the speed of searching. I have add a fulltext in text attribute of review table.

```
ALTER TABLE `yelp_db`.`review` DROP INDEX `text`, ADD FULLTEXT `text` (`text`);
```

And then add index to the ‘id’ attribute of business table and the ‘business\_id’ attribute of review table.

```
ALTER TABLE `yelp_db`.`review` DROP INDEX `business_id`, ADD INDEX `business_id`(`business_id`) USING BTREE;
```

```
ALTER TABLE `yelp_db`.`business` DROP INDEX `id`, ADD INDEX `id`(`id`) USINGBTREE;
```

## search.php

Connected successfully

Please input a keyword to search the business.

Keyword:

submit

Waiting for localhost...

## browsingPage3.php

Connected successfully

business_id	business_name
Nu_gZ_dTynmxLJsR1Mgeq	Motoreire
z9RSunhJgeXQWS4R36hkW	Buzz BBQ
jpYTsXFrojiBZ1kbWR21Zw	Four Peaks Grill & Tap
sR9hPrf4G-J-GLcl4yaiLw	Valley Bar
27_NSP0H7Z86viONPT_twg	Woodlot Restaurant
Y3aGrZR8Q8Bij8C2l7ofVdrg	The Vig McCormick
uGupeWjh0yIcCg8anM1PA	The Palazzo Las Vegas
q3PhW_SSeeEXHTO-zixLg	Yelp's Burger Bonanza @ Benjamins
9rCln1gnAqDVYU/XmlnY1w	5 Star BBQ & Grill
SG-117VX1LGByswntCKvrDQ	El Patenque
z519U_Hu4iVLwaCH6rc2w	Lo-Lo's Chicken and Waffles
K7IWdNuBChenEvI0NhGewg	Wicked Spoon
AmJ0GF1YCAap-vGYxv-OJQ	Rocky's Bar and Grill
cs1qJvnWUg_r8g020r_3PQ	Pizzeria Bianco
9NaZlhRPkxRJn1eA9zrTA	Saddle Mountain Brewing Company
NmOBonyz63TP9Nn2q4Z1rg	Smith & Wollensky
6PCoNYsEnpQVjssx4ysFw	Rock N Roll Tattoo
0ETipZSd6f7q_-o-7gutPg	Bradbury's
J8_AGg7P06s4B75MHBMEUQ	Tropical Smoothie Cafe
RVTxiiWQwneSlonJyYwwWg	teres : A Nail Bar
kZspuWnM0Y-Losvk2R0lA	Cous Cous Mediterranean Cafe
xtYHtTmuufCN2sUsQxBjA	Rio All Suites Hotel & Casino
JAmQCmc7UclNUfZknkjQ4A	High Roller
7PSZ4HTUM8M8Y9Yw3cc7Pg	Tipsy Coffee House
N-5i180DfZKH_Q5p_MoSx	Queso Good
-ubMwuskT2uSOGPSZpxaxA	Fleming's Prime Steakhouse - Peoria
rNL7RODe1D8SwU-U-FKg	The Katz Club Diner
ixV2SWdy7w8jzEAHp85qA	First Food & Bar
JFaTPxWQC14VmFnXLv7W1g	Melt Bar and Grilled

## detailPage.php(same php file as procedure)

Connected successfully

id	name	neighborhood	address	city	state	postal_code	latitude	longitude	stars	review_count	is_open
IZivKqtHyz4-ts8KsnvMrA	Kung Fu Tea	Chinatown	5030 Spring Mountain Rd	Las Vegas	NV	89146	36.127	-115.21	4.5	1041	1

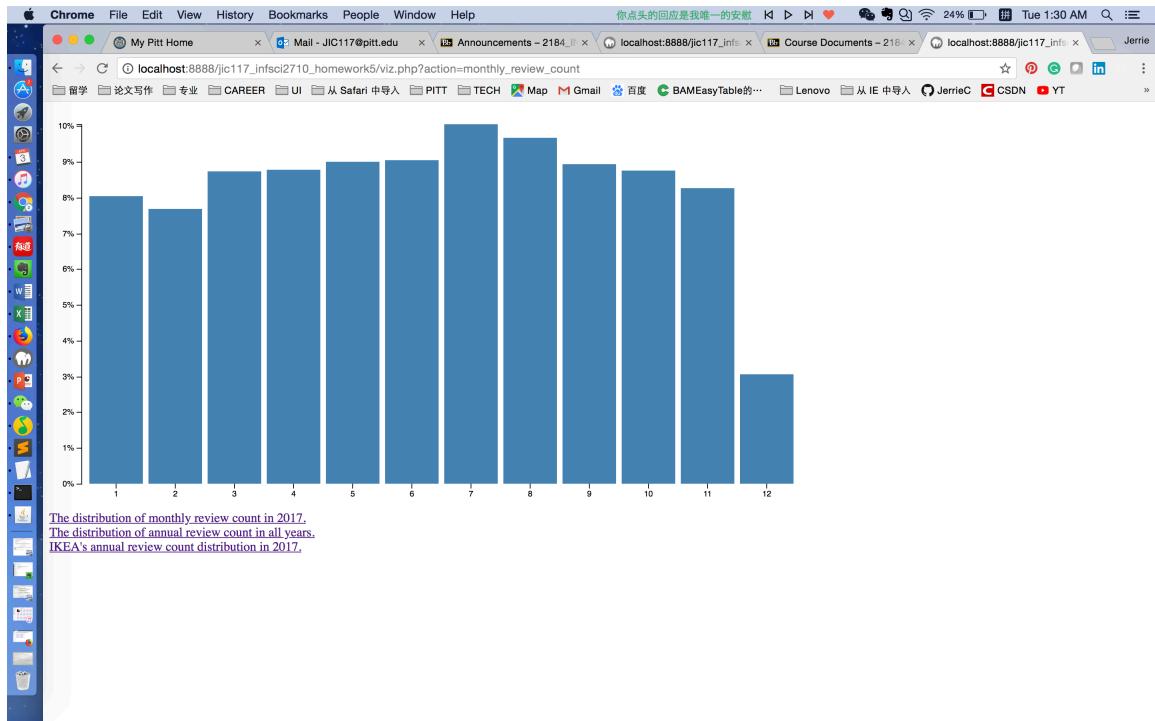
**Task 4 (25 points):** Write a data function (data.php) to generate a CSV for the given Bar Chart visualization (viz.php), so the user can inspect the Yelp data distribution of the following criteria:

1. The distribution of monthly review count in 2017.

```
select count(id) from review where year(date)=2017;  
SELECT month(date) as n, count(*)/1128518 as t from review WHERE year(date)=2017 group by n;
```

Showing rows 0 - 11 (12 total, Query took 31.0054 seconds.)

n	t
1	0.0804
2	0.0769
3	0.0874
4	0.0877
5	0.0901
6	0.0905
7	0.1005
8	0.0967
9	0.0893
10	0.0875
11	0.0826
12	0.0305



## 2. The distribution of annual review count in all years.

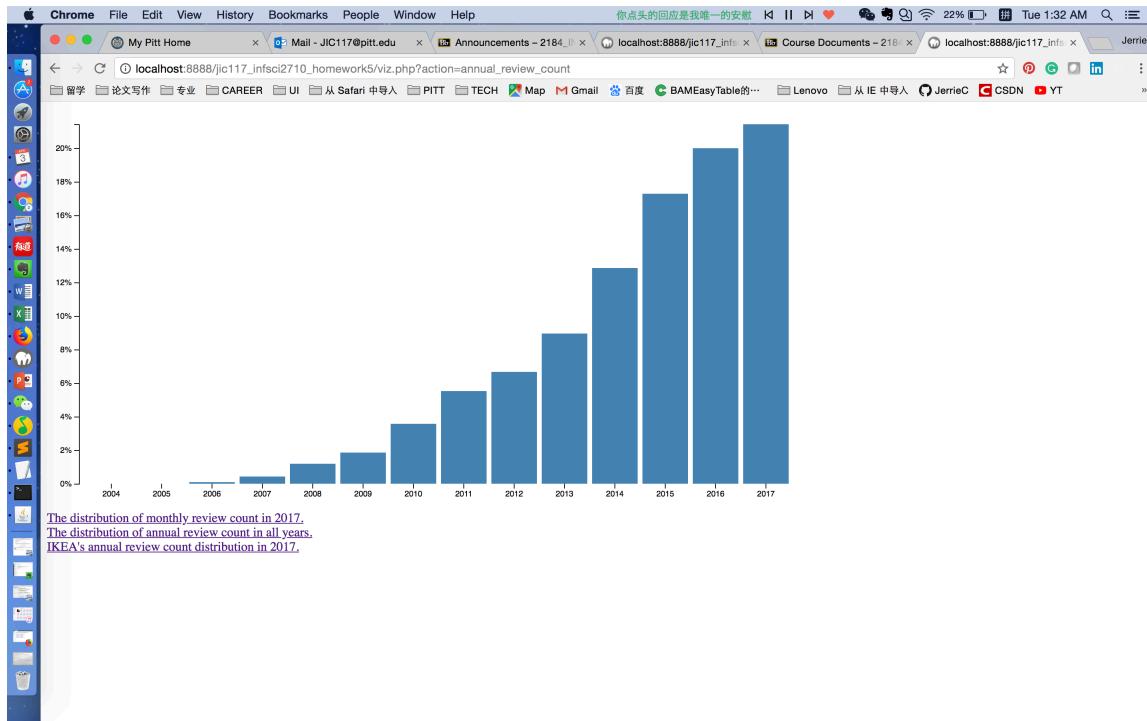
```
SELECT count(*) FROM review;
SELECT year(date) as n, count(*)/5261669 as t FROM review group by n;
```

Showing rows 0 - 13 (14 total, Query took 30.0315 seconds.)

year(date)	count(*)/5261669
2004	0.0000
2005	0.0002
2006	0.0011
2007	0.0044
2008	0.0117
2009	0.0187
2010	0.0356
2011	0.0553
2012	0.0666
2013	0.0898
2014	0.1289
2015	0.1732
2016	0.2001
2017	0.2145

+ Options

n	t
2004	0.0000
2005	0.0002
2006	0.0011
2007	0.0044
2008	0.0117
2009	0.0187
2010	0.0356
2011	0.0553
2012	0.0666
2013	0.0898
2014	0.1289
2015	0.1732
2016	0.2001
2017	0.2145

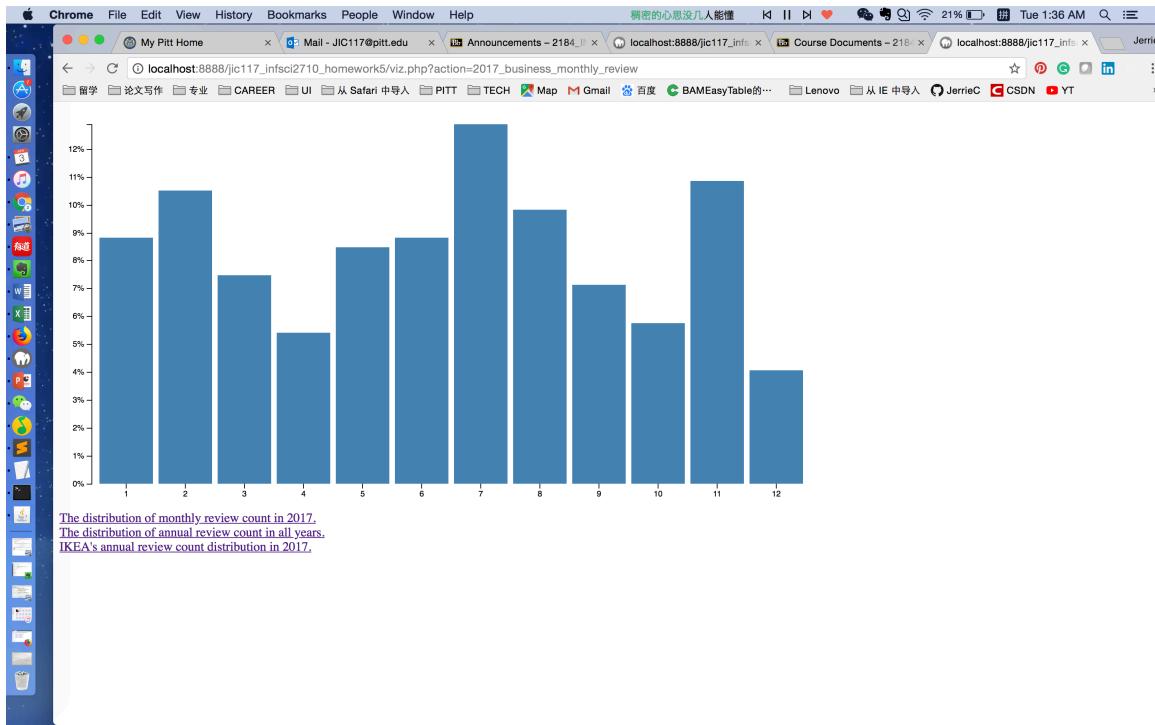


### 3. The distribution of monthly review count in 2017 of a certain business.

```
SELECT month(date) n, count(*)/(select count(*) from business b JOIN review r ON b.id = r.business_id WHERE b.name = business_name AND year(date) = 2017) t
FROM business b JOIN review r ON b.id = r.business_id WHERE b.name = business_name AND
year(date) = 2017 GROUP BY n;
CALL `2017_business_monthly_review` ('IKEA');
```

✓ Your SQL query has been executed successfully.  
12 rows affected by the last statement inside the procedure.

n	t
1	0.0881
2	0.1051
3	0.0746
4	0.0542
5	0.0847
6	0.0881
7	0.1288
8	0.0983
9	0.0712
10	0.0576
11	0.1085
12	0.0407



#### Hint:

1. Check <https://github.com/hymanct/infsci2710/tree/master/SampleCode/Visualization> for both viz.php and data.php files.
2. Check <https://bl.ocks.org/mbostock/3885304> for original example
3. You should modify the data.php
4. You should not modify the viz.php (besides line 70-77)

**Bonus Task 5 (25 Point):** Write a portal to integrate all the functions in task 1,2,3,4, so a user can visit the page and use all the proposed functions.

#### Note:

1. The code/script is expected to be Runnable. You will lose 20 points if your code is not runnable. I will use the following credential to test your code.

```
$servername = "localhost";      $username = "root";
$password = "mysql";          $database = "lab";
```

2. Collaboration on homework is permitted to an extent. Specifically, students are allowed to discuss the possible solutions to a problem and help each other with logic errors. However, handing your work to someone so that they may see a copy of your solution, or dictating code to a person on line-by-line basis is not within the spirit of the collaboration policy or the honor code of the university.
3. **I expect to see you demonstrate your works/solutions step by step.**