SQL Set Operations

SQL set operations

Allow us to query from more than one table at a time

SQL set operations

Sample Data:

amazon_best_sellers

year [PK] integer	title text	author text
2021	The Four Winds	Kristin Hannah
2020	A Promised Land	Barack Obama
2019	Where the Crawdads Sing	Delia Owens
2018	Becoming	Michelle Obama
2017	Wonder	R. J. Palacio

president_books

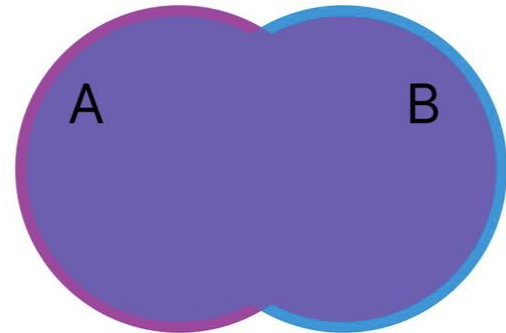|	id [PK] integer	title text	president text	year integer
1	1	Promises to Keep	Joe Biden	2008
2	2	A Promised Land	Barack Obama	2020
3	3	The Art of the Deal	Donald Trump	1987
4	4	Dreams From My Father	Barack Obama	1995
5	5	How to Get Rich	Donald Trump	2004
6	6	Decision Points	George W. Bush	2010
7	7	Between Hope and History	Bill Clinton	1996

SQL with Python

nucamp

UNION

In mathematics: the set of all elements in a collection
 - Written as A ∪ B

In SQL, the UNION operator is used to combine the result sets from two SELECT queries

Removes duplicates, returns combined result set

```
SELECT select_list_1
FROM table_1
UNION
SELECT select_list_2
FROM table_2;
```

UNION syntax

Both SELECT statements being combined must have:

Same **number** of columns

Same **order** of columns

Compatible **data types** for corresponding columns

```
SELECT select_list_1
FROM table_1
UNION
SELECT select_list_2
FROM table_2;
```

UNION - Example

Query all title, author, year combinations from amazon_best_sellers and president_books

```sql
SELECT title, president, year FROM president_books
UNION
SELECT title, author, year FROM amazon_best_sellers;
```

	title text	president text	year integer
1	Promises to Keep	Joe Biden	2008
2	How to Get Rich	Donald Trump	2004
3	Wonder	R. J. Palacio	2017
4	Between Hope and History	Bill Clinton	1996
5	Decision Points	George W. Bush	2010
6	The Four Winds	Kristin Hannah	2021
7	The Art of the Deal	Donald Trump	1987
8	Where the Crawdads Sing	Delia Owens	2019
9	Dreams From My Father	Barack Obama	1995
10	Becoming	Michelle Obama	2018
11	A Promised Land	Barack Obama	2020

SQL with Python

nucamp

Inaccurate

	title text	president text	year integer
1	Promises to Keep	Joe Biden	2008
2	How to Get Rich	Donald Trump	2004
3	Wonder	R. J. Palacio	2017
4	Between Hope and History	Bill Clinton	1996
5	Decision Points	George W. Bush	2010
6	The Four Winds	Kristin Hannah	2021
7	The Art of the Deal	Donald Trump	1987
8	Where the Crawdads Sing	Delia Owens	2019
9	Dreams From My Father	Barack Obama	1995
10	Becoming	Michelle Obama	2018
11	A Promised Land	Barack Obama	2020

```
SELECT title, president, year FROM president_books
UNION
SELECT title, author, year FROM amazon_best_sellers;
```

UNION - Example

Query all title, author, year combinations from amazon_best_sellers and president_books

```
SELECT title, president, year FROM president_books
UNION
SELECT title, author, year FROM amazon_best_sellers;
```

	title text	president text	year integer
1	Promises to Keep	Joe Biden	2008
2	How to Get Rich	Donald Trump	2004
3	Wonder	R. J. Palacio	2017
4	Between Hope and History	Bill Clinton	1996
5	Decision Points	George W. Bush	2010
6	The Four Winds	Kristin Hannah	2021
7	The Art of the Deal	Donald Trump	1987
8	Where the Crawdads Sing	Delia Owens	2019
9	Dreams From My Father	Barack Obama	1995
10	Becoming	Michelle Obama	2018
11	A Promised Land	Barack Obama	2020

UNION - Example

SQL with Python

nucamp

Query all title, author, year combinations from
amazon_best_sellers and president_books

```sql
SELECT title, president AS author, year FROM president_books
UNION
SELECT title, author, year FROM amazon_best_sellers;
```

	title text	author text	year integer
1	Promises to Keep	Joe Biden	2008
2	How to Get Rich	Donald Trump	2004
3	Wonder	R. J. Palacio	2017
4	Between Hope and History	Bill Clinton	1996
5	Decision Points	George W. Bush	2010
6	The Four Winds	Kristin Hannah	2021
7	The Art of the Deal	Donald Trump	1987
8	Where the Crawdads Sing	Delia Owens	2019
9	Dreams From My Father	Barack Obama	1995
10	Becoming	Michelle Obama	2018
11	A Promised Land	Barack Obama	2020

UNION - Example

Query all title, author, year combinations from amazon_best_sellers and president_books, **and allow duplicate rows**

```sql
SELECT title, president AS author, year FROM president_books
UNION ALL
SELECT title, author, year FROM amazon_best_sellers;
```

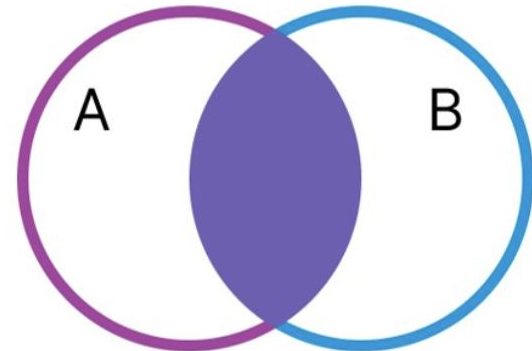|	title text	author text	year integer
1	Promises to Keep	Joe Biden	2008
2	A Promised Land	Barack Obama	2020
3	The Art of the Deal	Donald Trump	1987
4	Dreams From My Father	Barack Obama	1995
5	How to Get Rich	Donald Trump	2004
6	Decision Points	George W. Bush	2010
7	Between Hope and History	Bill Clinton	1996
8	Promises to Keep	Joe Biden	2008
9	A Promised Land	Barack Obama	2020
10	The Art of the Deal	Donald Trump	1987
11	Dreams From My Father	Barack Obama	1995
12	How to Get Rich	Donald Trump	2004

SQL with Python

nucamp

INTERSECT

In mathematics: Set intersection represents the set of all elements belonging to multiple sets
 - Written as A ∩ B

In SQL, INTERSECT operates on result sets of 2 queries

Combines result sets on top of one another, eliminates *unique* rows

```
SELECT select_list_1
FROM table_1
INTERSECT
SELECT select_list_2
FROM table_2;
```

INTERSECT - Example

Query all title, author, year combinations present in both
amazon_best_sellers and president_books

```sql
SELECT title, author, year FROM amazon_best_sellers
INTERSECT
SELECT title, president AS author, year FROM president_books;
```

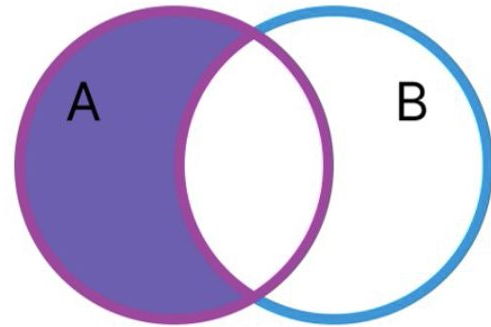|	title text	author text	year integer
1	A Promised Land	Barack Obama	2020

EXCEPT

EXCEPT implements set difference of two result sets

In mathematics: **Set difference** of A and B is the set of elements in A which are not in B
 - Written as A \ B

Returns first result set after eliminating any rows matching those of second result set

```
SELECT select_list_1
FROM table_1
EXCEPT
SELECT select_list_2
FROM table_2;
```

EXCEPT - Example

Query for amazon_best_sellers,
not including those written by presidents