Day 4

Challenge

Challenge

In the email system there was a problem with names where either the first name or the last name is more than 10 characters long.

Find these customers and output the list of these first and last names in all lower case.

Write a SQL query to find out!

Result

Data Output | Explain | Messages | Notifications

	lower text	lower text	lower text
1	william	satterfield	william.satterfield@sakilacustomer.org
2	christopher	greco	christopher.greco@sakilacustomer.org
3	henry	billingsley	henry.billingsley@sakilacustomer.org

Challenge

Challenge

In this challenge you have only the email address and the last name of the customers.

Data Output | Explain | Messages | Notifications

	email text	last_name text
1	MARY.SMITH@sakilacustomer.org	SMITH
2	PATRICIA.JOHNSON@sakilacustomer.org	JOHNSON
3	LINDA.WILLIAMS@sakilacustomer.org	WILLIAMS
4	BARBARA.JONES@sakilacustomer.org	JONES

You need to extract the first name from the email address and concatenate it with the last name. It should be in the form: "Last name, First name".

Write a SQL query to find out!

Result

Data Output | Explain | Messages

	?column? text
1	SMITH, MARY
2	JOHNSON, PATRICIA
3	WILLIAMS, LINDA

Challenge

Challenge

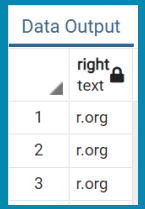
Extract the last 5 characters of the email address first.

The email address always ends with '.org'.

How can you extract just the dot '.' from the email address?

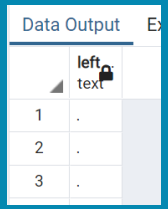
Write a SQL query to find out!

Result



Data Output

	right text
1	r.org
2	r.org
3	r.org



Data Output

	left text
1	.
2	.
3	.

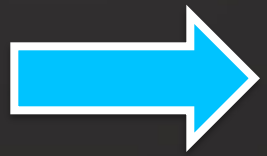
SUBSTRING

✓ **Used to EXTRACT a SUBSTRING from a string**

Data Output | Explain | Messages | Notifications

	email text
1	MARY.SMITH@sakilacustomer.org
2	PATRICIA.JOHNSON@sakilacustomer.org
3	LINDA.WILLIAMS@sakilacustomer.org

SUBSTRING



substring text
SMITH
JOHNSON
WILLIAMS

SYNTAX

Length,
How many characters?

```
SUBSTRING (string from start [for length] )
```

column / string
that we want to extract from

Position,
Where to start from?

SYNTAX

Length,
How many characters?

```
SUBSTRING (email from start [for length] )
```

column / string
that we want to extract from

Position,
Where to start from?

SYNTAX

Length,
How many characters?

```
SUBSTRING (email from 2 [for length] )
```

column / string
that we want to extract from

Position,
Where to start from?

SYNTAX

Length,
How many characters?

```
SUBSTRING (email from 2 for 3 )
```

column / string
that we want to extract from

Position,
Where to start from?

SYNTAX

Length,
How many characters?

```
SUBSTRING (email from 2 for 3 )
```

column / string
that w

Position,
t from?

Data Output | Explain | Messages | Notifications

	email text	substring text
1	MARY.SMITH@sakilacustomer.org	ARY
2	PATRICIA.JOHNSON@sakilacustomer.org	ATR
3	LINDA.WILLIAMS@sakilacustomer.org	IND

SYNTAX

```
SUBSTRING (email from 2 )
```

column / string
that we want to extract from

Position,
Where to start from?

Data Output | Explain | Messages | Notifications

	email text	substring text
1	MARY.SMITH@sakilacustomer.org	ARY.SMITH@sakilacustomer.org
2	PATRICIA.JOHNSON@sakilacustomer.org	ATRICIA.JOHNSON@sakilacustomer.org
3	LINDA.WILLIAMS@sakilacustomer.org	INDA.WILLIAMS@sakilacustomer.org

SYNTAX

Length,
How many characters?

```
SUBSTRING (email from POSITION ('.' in email) for 3 )
```

column / string
that w

Position,
t from?

Data Output | Explain | Messages | Notifications

	email text	substring text
1	MARY.SMITH@sakilacustomer.org	.SM
2	PATRICIA.JOHNSON@sakilacustomer.org	.JO
3	LINDA.WILLIAMS@sakilacustomer.org	.WI

SYNTAX

Length,
How many characters?

```
SUBSTRING (email from POSITION ('.' in email)+1 for 3 )
```

column / string
that w

Position,
t from?

Data Output | Explain | Messages | Notifications

	email text	substring text
1	MARY.SMITH@sakilacustomer.org	SMI
2	PATRICIA.JOHNSON@sakilacustomer.org	JOH
3	LINDA.WILLIAMS@sakilacustomer.org	WIL

SYNTAX

Length,
How many characters?

```
SUBSTRING (string from start [for length]
```

column / string
that we want to extract from

Position,
Where to start from?

Challenge

Challenge

You need to create an anonymized form of the email addresses in the following way:

1	M***.S***@sakilacustomer.org
2	P***.J***@sakilacustomer.org

In a second query create an anonymized form of the email addresses in the following way:

1	***Y.S***@sakilacustomer.org
2	***A.J***@sakilacustomer.org

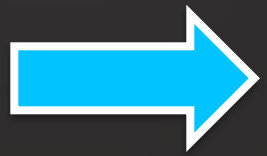
Write a SQL query to find out!

EXTRACT

✓ **Used to EXTRACT parts of timestamp/date**

rental_date timestamp with time zone
2005-05-24 23:54:33+02
2005-05-25 00:03:39+02
2005-05-25 00:04:41+02
2005-05-25 00:05:21+02

EXTRACT (day)



extract numeric
24
25
25
25

EXTRACT

✓ **Used to EXTRACT parts of timestamp/date**

rental_date timestamp with time zone
2005-05-24 23:54:33+02
2005-05-25 00:03:39+02
2005-05-25 00:04:41+02
2005-05-25 00:05:21+02

EXTRACT (seconds)

extract numeric
33.000000
39.000000
41.000000
21.000000

Date/time types

`date`	Just date without time	`'2022-11-28'`
`time (with/without time zone)`	Just time without date	`'01:02:03.678'`
`timestamp (with/without time zone)`	Date and time	`'2022-11-28 01:02:03.678+02'`
`intervals`	Time interval	`'3 days 01:02:03.678'`

SYNTAX

```
EXTRACT (field from date/time/interval)
```

Part of date/time

Date/time
that we want to extract from

EXTRACT

Usually singular

Useful when creating new tables

Field	Extract from timestamp/date
CENTURY	century
DAY	day of month (1-31)
DECADE	decade that is year divided by 10
DOW	day of week Sunday (0) to Saturday (6)
DOY	day of year that ranges from 1 to 366
EPOCH	number of seconds since 1970-01-01 00:00:00 UTC
HOUR	hour (0-23)
ISODOW	day of week based on ISO 8601 Monday (1) to Sunday (7)
ISOYEAR	ISO 8601 week number of year
MICROSECONDS	seconds field, including fractional parts, multiplied by 1000000
MILLENNIUM	millennium
MILLISECONDS	seconds field, including fractional parts, multiplied by 1000
MINUTE	minute (0-59)
MONTH	month (1-12)
QUARTER	quarter of year
SECOND	second
TIMEZONE	timezone offset from UTC, measured in seconds
TIMEZONE_HOUR	hour component of time zone offset
TIMEZONE_MINUTE	minute component of time zone offset
WEEK	number of ISO 8601 week-numbering week of year
YEAR	year

EXTRACT

Usually singular

Useful when creating new tables

Field	Extract from timestamp/date
CENTURY	century
DAY	day of month (1-31)
DECADE	decade that is year divided by 10
DOW	day of week Sunday (0) to Saturday (6)
DOY	day of year that ranges from 1 to 366
EPOCH	number of seconds since 1970-01-01 00:00:00 UTC
HOUR	hour (0-23)
ISODOW	day of week based on ISO 8601 Monday (1) to Sunday (7)
ISOYEAR	ISO 8601 week number of year
MICROSECONDS	seconds field, including fractional parts, multiplied by 1000000
MILLENNIUM	millennium
MILLISECONDS	seconds field, including fractional parts, multiplied by 1000
MINUTE	minute (0-59)
MONTH	month (1-12)
QUARTER	quarter of year
SECOND	second
TIMEZONE	timezone offset from UTC, measured in seconds
TIMEZONE_HOUR	hour component of time zone offset
TIMEZONE_MINUTE	minute component of time zone offset
WEEK	number of ISO 8601 week-numbering week of year
YEAR	year

Challenge

You need to analyze the payments and find out the following:

- What's the month with the highest total payment amount?
- What's the day of week with the highest total payment amount? (0 is Sunday)
- What's the highest amount one customer has spent in a week?

Write a SQL query to find out!

Result

	month numeric	total_payment_amount numeric
1	4	28327.02
2	3	23886.56

	day_of_week numeric	total_payment_amount numeric
1	4	12796.08
2	1	12132.12

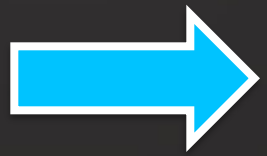
	week numeric	customer_id smallint	total_payment_amount numeric
1	18	459	73.88
2	12	21	72.86
3	18	2	65.88

TO_CHAR

- ✓ **Used to get custom formats timestamp/date/numbers**

rental_date timestamp with time zone
2005-05-24 23:54:33+02
2005-05-25 00:03:39+02
2005-05-25 00:04:41+02
2005-05-25 00:05:21+02

TO_CHAR (YYYY-MM)



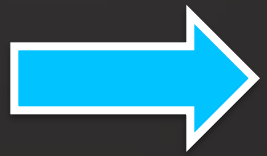
	to_char text
1	2005-05
2	2005-05
3	2005-05
4	2005-05

TO_CHAR

- ✓ **Used to get custom formats timestamp/date/numbers**

rental_date timestamp with time zone
2005-05-24 23:54:33+02
2005-05-25 00:03:39+02
2005-05-25 00:04:41+02
2005-05-25 00:05:21+02

TO_CHAR (Month)



	to_char text
1	May
2	May
3	May
4	May

SYNTAX

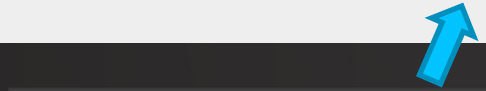
```
TO_CHAR (date/time/interval, format)
```

date/time/interval/number

Format

SYNTAX

```
TO_CHAR (rental_date, format)
```



SYNTAX

```
TO_CHAR (rental_date, 'MM-YYYY')
```

date/time/interval/number

Format

ym text
05-2020
03-2020
04-2020

Challenge

You need to sum payments and group in the following formats:

total_amount numeric	day text
62.86	Fri, 24/01/2020
70.81	Fri, 14/02/2020

	total_amount numeric	day text
1	746.62	May, 2020
2	4824.43	Jan, 2020

	total_amount numeric	day text
1	537.14	Thu, 02:44
2	59.90	Wed, 10:06

Write a SQL query to find out!

Result

	month numeric	total_payment_amount numeric
1	4	28327.02
2	3	23886.56

	day_of_week numeric	total_payment_amount numeric
1	4	12796.08
2	1	12132.12

	week numeric	customer_id smallint	total_payment_amount numeric
1	18	459	73.88
2	12	21	72.86
3	18	2	65.88

Challenge

You need to create a list for the suppcity team of all rental durations of customer with customer_id 35.

Also you need to find out for the suppcity team
which customer has the longest average rental duration?

Write a SQL query to find out!

Result

	customer_id smallint	rental_duration interval
1	35	4 days 20:59:00
2	35	8 days 18:10:00
3	35	5 days 01:12:00

	customer_id smallint	avg interval
1	315	6 days 14:13:22.5
2	187	5 days 34:58:38.571428
3	321	5 days 32:56:32.727273