

CompTIA DataX Exam

Question & Answers Demo

# Version: 4.0

Question: 1

SIMULATION

A client has gathered weather data on which regions have high temperatures. The client would like a visualization to gain a better understanding of the data.

INSTRUCTIONS

Part 1

Review the charts provided and use the drop-down menu to select the most appropriate way to standardize the data.

Part 2

Answer the questions to determine how to create one data set.

Part 3

Select the most appropriate visualization based on the data set that represents what the client is looking for.

If at any time you would like to bring back the initial state of the simulation, please click the Reset All button.

Part 1	Part 2	Part 3						
Standardi	ze data							
Select tab	le	~)	+	Table 1				
Table 2				City		State	Zip code	Region
		2		Orla	ando	FL	32802	South
				New	York	NY	10001	North
				De	enver	СО	80014	West
				New Orl	eans	LA	7003	Central
				Richn	nond	VA	23173	East
				Table 2	Zin-s	T. abo	opporative	o Scolo
				Couth	210 0	000	emperatur	
				North	10	001	5	0 °F
				West	80	001	3	
				Central	00	NaN	6	2 0
				East	23	173	5	0 °C

rarez rarez		
Standardize data		
Select table	Table 1	
Table 1	City State Zip code Reg	aion
Variable:	Orlando El 32802 So	outh
Select variable to standardize	New York NY 10001 No	orth
City		Voct
Zip code		vest
кедюл	New Orleans LA 7003 Cen	ntral
B - 11	THE REPORT OF THE PARTY OF THE	
Select action to take   Remove	Richmond VA 23173 E	East
Action: Select action to take V Remove Correct	Richmond VA 23173 E	East
Action: Select action to take V Remove Correct	Richmond VA 23173 E Table 2 Region Zip code Temperature So	eale
Select action to take   Remove Correct	Richmond VA 23173 E Table 2 Region Zip code Temperature So South 32802 50	cale °F
Select action to take  Remove Correct	Richmond VA 23173 E	cale °F
Select action to take   Remove Correct	Richmond       VA       23173       E         Table 2       Region       Zip code       Temperature       So         South       32802       50       50         North       10001       68       68         West       80014       30	cale °F °F
Select action to take  Remove Correct	Richmond       VA       23173       E         Table 2       Region       Zip code       Temperature       So         South       32802       50       50         North       10001       68       68         West       80014       30       62	cale °F °F
Select action to take   Remove Correct	Richmond       VA       23173       E         Table 2       Region       Zip code       Temperature       So         South       32802       50       50         North       10001       68       68         West       80014       30       62         South       32170       50       50	cale °F °F °F

Part 1 Part 2	Part 3						
Standardize data							
Select table	~	+	Table 1	3			
Table 1		×	City		State	Zip code	Region
/ariable:			Orl	ando	FI	32802	South
State	¥	)	Now	Vork	NV	10001	North
Select action to take	~	1	De	nor	0	80014	Mort
Remove			Now Orl	loanc		7002	Control
Correct	~		New On	learis		/003	Central
			Table 2				
			Region	Zip c	ode T	emperatur	e Scale
			South	32	802	5	0 °F
			North	10	001	6	8 °F
			West	80	014	3	0 °F
			Central		NaN	6	2 °F
			East	23	173	5	0 °C

Part 1 Part 2 Part 3	
Standardize data	
Select table ~	+ Table 1
Table 1	City State Zip code Region
/ariable:	Orlando FL 32802 South
Action:	New York NY 10001 North
Select action to take	Denver CO 80014 West
Remove	New Orleans LA 7003 Central
	Richmond VA 23173 East
	Table 2
	Region Zip code Temperature Scale
	South 32802 50 °F
	North 10001 68 °F
	West 80014 30 °F
	Central NaN 62 °F
	East 23173 50 °C

Part 1 Part 2 Part 3				
Standardize data				
Select table ~	+ Table 1			
able 1	City	State	Zip code	Region
ariable:	Orlan	ndo FL	32802	South
lip code N	New Y	ork NY	10001	North
Select action to take	Den	ver CO	80014	West
Remove	New Orlea	ans LA	7003	Central
032802 010001 080014 023173	Richmo	ond VA	23173	East
	Table 2			
	Region Z	lip code T	emperatur	e Scale
	South	32802	5	0 °F
	North	10001	6	8 °F
	West	80014	3	0 °F
	Central	NaN	6	2 °F
	East	23173	5	0 °C

Part 1 Part 2	Part 3						
Standardize data							
Select table	~	+	Table 1	1			
Table 1		$\mathbf{x}$	Citiy		State	Zip code	Region
/ariable:			Orla	ando	FL	32802	South
kegion			New	York	NY	10001	North
Select action to take	Ý	]	De	enver	со	80014	West
Remove Correct			New Orl	eans	LA	7003	Central
	t O East	1	Richn	nond	VA	23173	East
			Table 2				
			Region	Zip o	ode T	emperatur	e Scale
			South	32	802	5	0 °F
			North	10	001	6	8 °F
			West	80	014	3	0 °F
			Central		NaN	6	2 °F
			East	23	173	5	0 °C

art 1 Part 2 Part 3				
tandardize data				
elect table 🛛 👻 🕂	Table 1			
able 2 🛞	City	State	Zip code	Region
ariable:	Orland	o FL	32802	South
elect variable to standardize	New Yor	k NY	10001	North
legion	Denve	er CO	80014	West
ction:	New Orlean	s LA	7003	Centra
select action to take	Richmon	d VA	23173	East
	Table 2			
	Region Zip	o code T	emperatur	e Scale
	South 3	32802	5	0 °F
	North :	10001	6	0 00
				0 1
	West 8	30014	3	0 °F
	West 8 Central	30014 NaN	3	0 °F 2 °F

tandardize data				
elect table 💙 🕂				
	Table 1	and the second s		lious
able 2	City	State	Zip code	Region
Zip code	Orland	o FL	32802	South
ction:	New Yor	k NY	10001	North
Select action to take	Denve	er CO	80014	West
kemove				-
Correct	New Orlear	is LA	/003	Central
Correct O NaN O 23173 O 32802 O 10001 O 80014	New Orlear Richmon	is LA d VA	23173	East
Correct O NaN O 23173 O 32802 O 10001 O 80014	New Orlear Richmon	is LA d VA	23173	East
Correct O NaN O 23173 O 32802 O 10001 0 80014	New Orlear Richmon Table 2 Region Zig	d VA	23173 emperatur	East
orrect NaN 023173 032802 010001 80014	New Orlear Richmon Table 2 Region Zig South	d VA	7003 23173 emperatur 5	Centra East
orrect NaN 023173 032802 010001 80014	New Orlear Richmon Table 2 Region Zip South 3 North	o code 1 32802 10001	7003 23173 emperatur 5 6	Centra East e Scale 0 °F 8 °F
orrect NaN 023173 032802 010001 80014	New Orlear Richmon	o code 1 32802 10001 30014	7003 23173 emperatur 5 6 3	Centra East Centra East Cale O PF 8 PF 0 PF
Correct O NaN O 23173 O 32802 O 10001 0 80014	New Orlear Richmon	o code 1 32802 10001 30014 NaN	7003 23173 emperatur 5 6 3 6	Centra East 0 °F 8 °F 0 °F 2 °F

Part 1	Part 2	Part 3						
Standardi	ze data							
Select tab	ole	~)	+	Table 1				
Table 2			×	City	ù	State	Zip code	Region
ariable:			5	Orl	ando	FL	32802	South
ction:		~	,	New	York	NY	10001	North
Select actio	on to take	~	2	De	enver	со	80014	West
Remove Correct				New Orl	eans	LA	7003	Central
O South C	North OWes	t O East		Richn	nond	VA	23173	East
				Table 2				
				Region	Zip d	ode T	emperatur	e Scale
				South	32	802	5	0 °F
				North	10	001	6	8 °F
				West	80	014	3	0 °F
				Central		NaN	6	2 °F
				East	23	173	5	0 °C

Part 1	Part 2	Part 3						
Standardize	e data							
Select table	е	~	+	Table 1				
able 2			X	City	2 	State	Zip code	Region
ariable:	I and		Y	Orla	ando	FL	32802	South
emperature, ction:	/scale	· ·	<i>)</i> //	New	York	NY	10001	North
Select action	to take	~	2	De	enver	со	80014	West
Remove Correct				New Orl	eans	LA	7003	Central
062°F 030	0°F ○ 50°C (	⊃68°F		Richn	hond	VA	23173	Fast
⊃ 50°F				rderin				2.00
⊃ 50°F				Table 2				
⊃ 50°F				Table 2 Region	Zip c	ode T	emperatur	e Scale
) 50°F				Table 2 Region South	Zip o 32	ode T 802	emperatur 5	e Scale
⊃ 50°F				Table 2 Region South North	Zip c 32 10	ode T 802 001	emperatur 5	e Scale 0 °F 8 °F
⊃ 50°F				Table 2 Region South North West	Zip c 32 10 80	ode T 802 001 014	emperatur 5 6 3	e Scale 0 °F 8 °F 0 °F
⊖ 50°F				Table 2 Region South North West Central	Zip c 32 10 80	ode T 802 001 014 NaN	emperatur 5 6 3 6	e Scale 0 °F 8 °F 0 °F 2 °F

Merge data         Select the most appropriate method to use when combining these two tables:         O Data matching       O Filter         O Union       O Deduplication         Select the most appropriate variable to use when joining these sets of data:	Table 1 City Orla New De New Orla	ando York enver eans	State FL NY CO	Zip code 32802 10001 80014	Region South North
Select the <b>most</b> appropriate method to use when combining these two tables: O Data matching O Filter O Union O Deduplication Select the <b>most</b> appropriate variable to use when joining these sets of data:	Table 1 City Orla New De New Orla	ando York enver eans	State FL NY CO	Zip code 32802 10001 80014	Region South North
O Data matching O Filter O Union O Deduplication Select the <b>most</b> appropriate variable to use when joining these sets of data:	City Orla New De New Orla	ando York enver eans	State FL NY CO	Zip code 32802 10001 80014	Region South North
O Data matching O Filter O Union O Deduplication Select the <b>most</b> appropriate variable to use when joining these sets of data:	Orla New De New Orle	ando York enver eans	FL NY CO	32802 10001 80014	South North
O Union O Deduplication Select the <b>most</b> appropriate variable to use when joining these sets of data:	New De New Orle	York nver eans	NY CO	10001 80014	North
Select the <b>most</b> appropriate variable to use when joining these sets of data:	De New Orle	enver eans	СО	80014	Most
use when joining these sets of data:	New Orl	eans			west
use when joining these sets of tidta.	All and the second second		LA	7003	Central
	Richm	nond	VA	23173	East
O Zip code					
	Table 2				
	Region	Zipo	ode T	emperatur	e Scale
	South	32	802	5	0 °F
	North	10	001	6	8 °F
	West	80	014	3	0 °F
	Central	)	NaN	6	2 °F
	East	23	173	5	0 °C

Part :	L Par	t 2 Part 3				
Visual	ization					
	Sele	ect the <b>most</b> ap a set which rep	opropriat resents v	e visualizat what the cli	ion based on the ent is looking fo	e r:
0	· 》	°	 			Action and a second sec
		。 便	X	0	en and	
	Region	City	State	Zip code	Temperature	Scale
	South	Orlando	FL	32802	50	٩F
	North	New York	NY	10001	68	٩F
	West	Denver	CO	80014	30	٩F
	Central	New Orleans	LA	7003		
	East	Richmond	VA	23173	50	°C
	Central			NaN	62	٩F













Temperature

Answer: See explanation below.

Explanation:

Explanation:

Part 1

Select Table 2. Table 2 contains mixed temperature scales (°F and °C) that must be standardized before visualization.

Variable: Temperature/scale

Action: Correct

Value to correct: 50 °C

Part 1	Part 2	Part 3						
Standardi	ze data							
Select tab Table 2 Variable: Temperatur Action: Select action Remove Correct	ne/scale		+	Table 1 City Orla New De New Orl	ando York enver leans	State FL NY CO LA	Zip code 32802 10001 80014 7003	Region South North West Central
062°F 0	30°F 250°C (	⊖ <b>68°</b> F		Richn	nond	VA	23173	East
				Table 2 Region	Zip o	ode T	emperatur	e Scale
				South	32	802	5	0 °F
				North	10	001	6	8 °F
				West	80	014	3	0 °F
				Central		NaN	6	2 °F
				East	23	173	5	0 °C

#### Part 2

Method: Data matching

Join variable: Zip code

You need to merge the two tables by aligning matching records, which is a data-matching (join) operation, and ZIP code is the shared, uniquely identifying field linking each region's weather reading to its city.

Part 1 Part 2	Part 3						
Merge data							
Select the <b>most</b> appro- use when combining t	opriate metho hese two tabl	d to es:	Table 1		Chata	Zin code	Docion
Data matching	Filter		Orla	ando	FL	32802	South
	Deduplication		New	York	NY	10001	North
Select the <b>most</b> appro	Select the <b>most</b> appropriate variable to				LA	7003	Central
<ul> <li>Region</li> <li>Zip code</li> </ul>			Richn	nond	VA	23173	East
			Table 2				
			Region	Zip c	ode T	emperatur	e Scale
			South	32	802	5	0 °F
			North	10	001	6	8 °F
			West	80	014	3	0 °F
			Central		NaN	6	2 °F
			East	23	173	5	0 °C

#### Part 3

Choose the choropleth map (the first option).

A choropleth map best shows geographic variation in temperature by coloring each state (or region) according to its recorded value. This lets the client immediately see where the highest and lowest temperatures occur across the U.S. without distracting elements like bubble size or combined chart axes.

Part 1	Par	t 2 Part 3					
Visualiz	zation						
	Sele	ect the <b>most</b> ap a set which repr	propriat resents v	e visualizat what the cli	ion based on the ent is looking fo	e r:	
			 			lat fast	
		。 便	X	0	NE BEI Trease		
	Region	City	State	Zip code	Temperature	Scale	
	South	Orlando	FL	32802	50	٩	
	North	New York	NY	10001	68	٩F	
	West	Denver	CO	80014	30	°F	
	Central	New Orleans	LA	7003			
	East	Richmond	VA	23173	50	°C	
	Central			NaN	62	٩F	

Question: 2	
SIMULATION	

A data scientist needs to determine whether product sales are impacted by other contributing factors. The client has provided the data scientist with sales and other variables in the data set.

The data scientist decides to test potential models that include other information.

### INSTRUCTIONS

Part 1

Use the information provided in the table to select the appropriate regression model.

Part 2

Review the summary output and variable table to determine which variable is statistically significant.

If at any time you would like to bring back the initial state of the simulation, please click the Reset All button.













	Var 1 Sales	Var 2 ROI	Var 3	Var	
Time	(in millions)	(% of overall)	Inventory cost	oper	
1	326.311584	16%	58		
2	507.9584031	8%	57		
3	232.5685962	5%	53		
4	117.3342091	7%	50		
5	242.866515	7%	60		
6	359.6300247	14%	50		
7	119.384542	19%	56		
8	372.064584	5%	56		
9	320.0212452	18%	51		
	Image: Strategy of	Which of the foll include in the ne O Var 5 Initial inve	lowing additional va ew model? estment O Var 4	riable Net op	

Summary output							
Regression statistics					Coefficients	Standard error	ts
Multiple R	0.9	99978259	Inter	rcept	30.24229003	9.306229821	3
R square	0.9	99956518	Var 2 (% of ove	ROI erall)	50.72139711	13.14967361	:
Adjusted R square	0.9	99913036	Inventory	Var 3 cost	-0.315571292	2.013342425	
Standard error	1.1	00286825	Var 4 Net operations cost		9.854244454	0.049842563	
Observations		9	Var 5 I investr	nitial ment	-0.268287655	0.103591751	
	df	SS		MS		F	
Regression	4		111363.9712		27840.992	8 2299	7.09
Residual	4		4.842524393		1.21063109	8	
Total	8		111368.8137				

Answer: See explanation below.

Explanation:

Part 1

Linear regression.

Of the four models, linear regression has the highest  $R^2$  (0.8), indicating it explains the greatest proportion of variance in sales.

Giver	Part 2	ues, wh	ich of the following regres	sion models <b>best</b> fits the re	elationship between the
0	Diesr		0		
	Ridge regre	ssion	Quantile regression	Linear regression	Lasso regression
	R <sup>2</sup> 0.5		R <sup>2</sup> 0.6	R <sup>2</sup> 0.8	R <sup>2</sup> 0.62
	1	Time	Var 1 Sales (in millions)	Var 2 ROI (% of overall)	R <sup>2</sup> Value
		1	3.118026935	6%	
		2	4.823728572	11%	
		3	7.149131157	18%	
		4	2.173859679	5%	
		5	3.519662597	9%	
		6	5.98246748	12%	
		7	8.495414141	14%	
		8	3.678906129	7%	
		9	3.539605808	6%	

#### Part 2

Var 4 – Net operations cost.

Net operations cost has a p-value of essentially 0 (far below 0.05), indicating it is the only additional predictor statistically significant in explaining sales. Neither inventory cost ( $p\approx0.90$ ) nor initial investment ( $p\approx0.23$ ) reach significance.

Time	Var 1 Sales (in millions)	Var 2 ROI (% of overall)	Var 3 Inventory cost	Var 4 Net operations cost
1	326.311584	16%	58	32
2	507.9584031	8%	57	50
3	232.5685962	5%	53	23
4	117.3342091	7%	50	11
5	242.866515	7%	60	24
6	359.6300247	14%	50	35
7	119.384542	19%	56	11
8	372.064584	5%	56	37
9	320.0212452	18%	51	31
	Annual An	Which of the follo include in the new O Var 5 Initial inves	wing additional va w model? stment Var 4	riables should the dat Net operations cost

## **Question: 3**

A data scientist is building an inferential model with a single predictor variable. A scatter plot of the independent variable against the real-number dependent variable shows a strong relationship between them. The predictor variable is normally distributed with very few outliers. Which of the following algorithms is the best fit for this model, given the data scientist wants the model to be easily interpreted?

A. A logistic regression

B. An exponential regression

C. A linear regression

D. A probit regression

Answer: C

Explanation:

Question: 4

A data scientist wants to evaluate the performance of various nonlinear models. Which of the following is best suited for this task?

A. AIC B. Chi-squared test C. MCC D. ANOVA

Answer: A

Explanation:

# Question: 5

Which of the following is the layer that is responsible for the depth in deep learning?

- A. Convolution
- B. Dropout
- C. Pooling
- D. Hidden

Answer: D

Explanation: