

A wide-angle photograph of a long, straight asphalt road that stretches from the foreground into the distance, vanishing at the horizon. The road is flanked by rolling hills covered in sparse, dry vegetation and patches of snow. The sky is overcast with heavy, grey clouds. The overall mood is desolate and expansive.

// .XPRESSION



Technology increases  
productivity... **Right?**

431,000

The number of people injured from  
distracted driving in 2014.



# DISTRACTIONS

What features in a scene are correlated with distraction?





# DISTRACTIONS

Who cares?



# DISTRACTIONS

Who cares?

Market:  
**PHILIPS**

BRITISH AIRWAYS 

**VOLVO**

**TERADATA**



# Tools:

TERADATA

TERADATA.  
ASTER



kaggle™



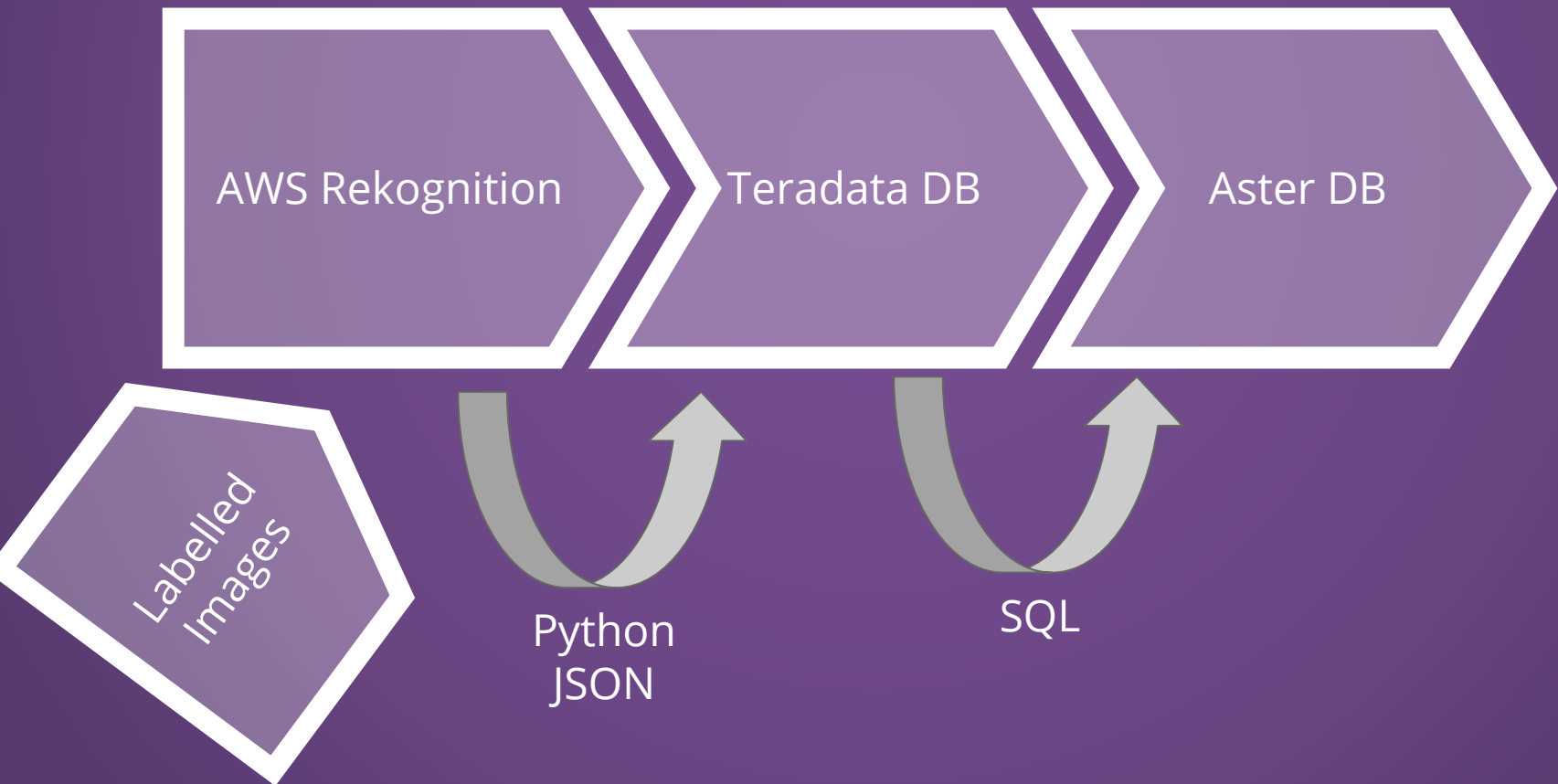
# THE PLAN

AWS Rekognition

Teradata DB

Aster DB

# THE PLAN



# GETTING DATA

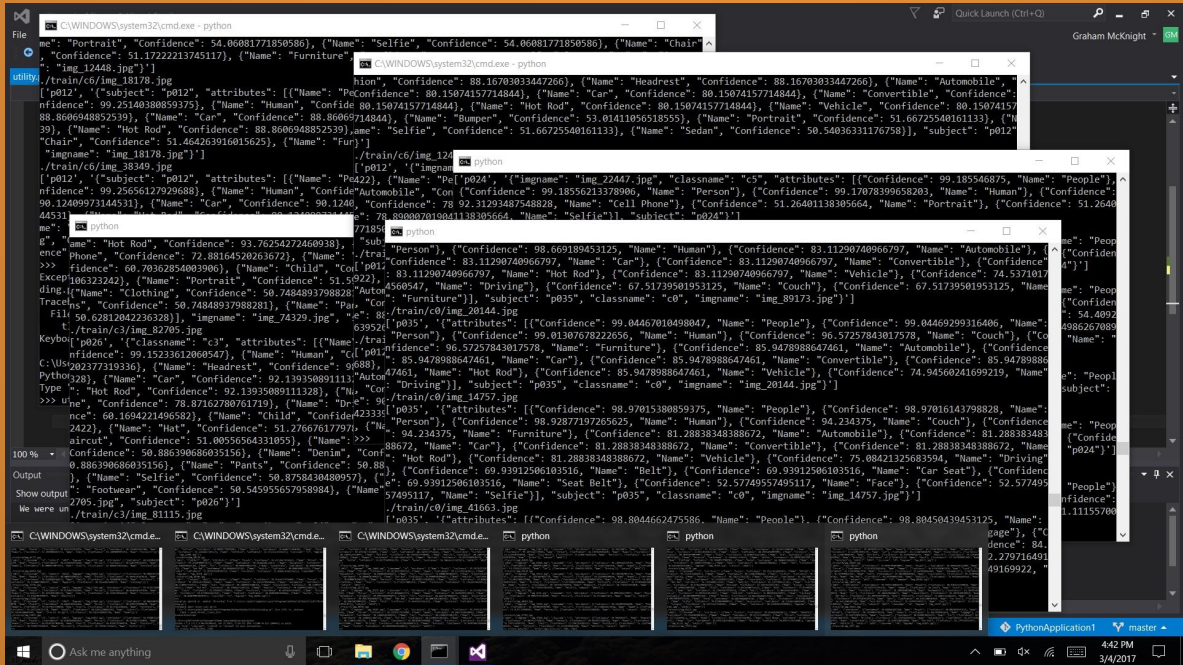
Labelled Dataset from Kaggle  
(StateFarm Competition)  
20,000 Images



# AWS Rekognition

Uploaded using Python scripts

More Python scripts to get several requests simultaneously





## Teradata Database

First converted JSON from Rekognition to CSV (Python)

Reformatted to sort by Rekognition tags

Removed least common tags (Potted plant, Whale, Lingerie...)

Uploaded data to Teradata DB

# Aster Analytics

At first we saw a lot of error messages...



To log on to this remote computer, you must be granted the Allow log on through Terminal Services right. By default, members of the Remote Desktop Users group have this right. If you are not a member of the Remote Desktop Users group or another group that has this right, or if the Remote Desktop Users group does not have this right, you must be granted this right manually.

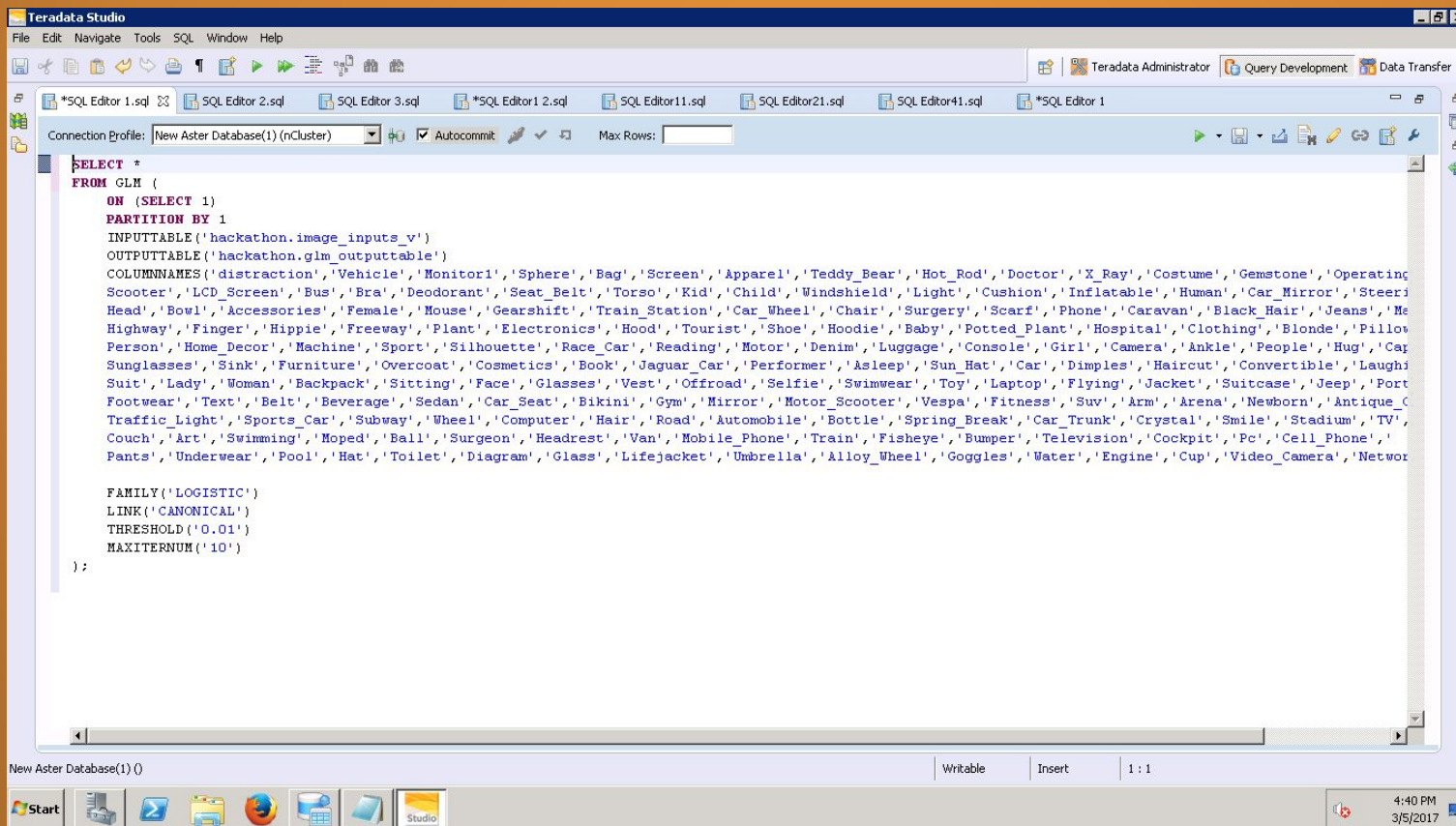
OK



Windows Server<sup>™</sup> 2008 R2  
Datacenter

# Aster Analytics

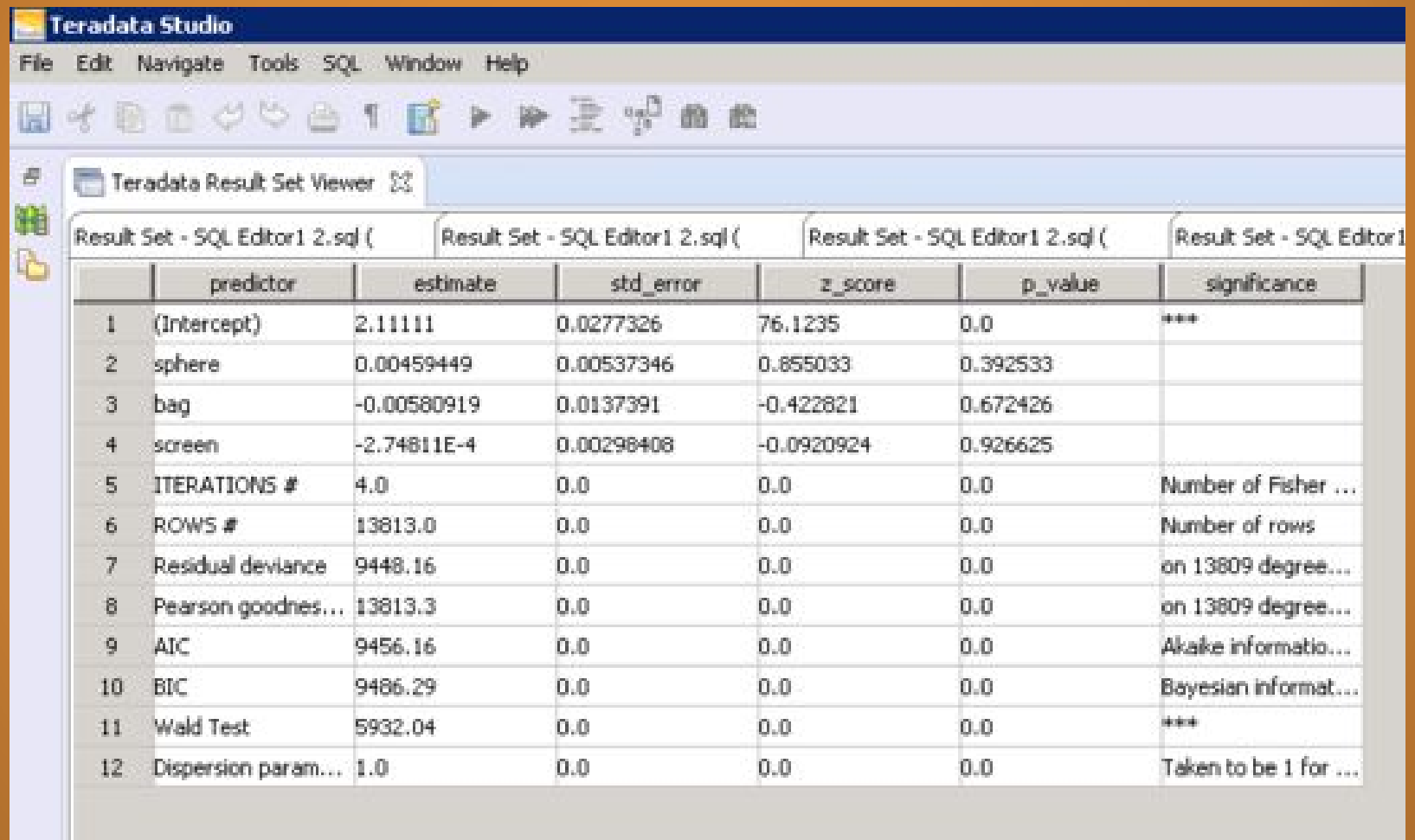
But ended up getting GLM working!



The screenshot shows the Teradata Studio interface with a SQL Editor window open. The query being executed is a GLM (Generalized Linear Model) query. The query starts with a SELECT statement, followed by a FROM clause specifying the GLM model. The model is defined by a SELECT statement with a PARTITION BY clause, an INPUTTABLE, and an OUTPUTTABLE. The COLUMNS clause lists a large number of features, including 'distraction', 'Vehicle', 'Monitor', 'Sphere', 'Bag', 'Screen', 'Apparel', 'Teddy\_Bear', 'Hot\_Rod', 'Doctor', 'X-Ray', 'Costume', 'Gemstone', 'Operating\_Scooter', 'LCD\_Screen', 'Bus', 'Bra', 'Deodorant', 'Seat\_Belt', 'Torso', 'Kid', 'Child', 'Windshield', 'Light', 'Cushion', 'Inflatable', 'Human', 'Car\_Mirror', 'Steering\_Head', 'Bowl', 'Accessories', 'Female', 'Mouse', 'Gearshift', 'Train\_Station', 'Car\_Wheel', 'Chair', 'Surgery', 'Scarf', 'Phone', 'Caravan', 'Black\_Hair', 'Jeans', 'Me\_Highway', 'Finger', 'Hippie', 'Freeway', 'Plant', 'Electronics', 'Hood', 'Tourist', 'Shoe', 'Hoodie', 'Baby', 'Potted\_Plant', 'Hospital', 'Clothing', 'Blonde', 'Pillow\_Person', 'Home\_Decor', 'Machine', 'Sport', 'Silhouette', 'Race\_Car', 'Reading', 'Motor', 'Denim', 'Luggage', 'Console', 'Girl', 'Camera', 'Ankle', 'People', 'Hug', 'Cap', 'Sunglasses', 'Sink', 'Furniture', 'Overcoat', 'Cosmetics', 'Book', 'Jaguar\_Car', 'Performer', 'Asleep', 'Sun\_Hat', 'Car', 'Dimples', 'Haircut', 'Convertible', 'Laugh\_Suit', 'Lady', 'Woman', 'Backpack', 'Sitting', 'Face', 'Glasses', 'Vest', 'Offroad', 'Selfie', 'Swinwear', 'Toy', 'Laptop', 'Flying', 'Jacket', 'Suitcase', 'Jeep', 'Port\_Footwear', 'Text', 'Belt', 'Beverage', 'Sedan', 'Car\_Seat', 'Bikini', 'Gym', 'Mirror', 'Motor\_Scooter', 'Vespa', 'Fitness', 'Suv', 'Arm', 'Arena', 'Newborn', 'Antique\_C Traffic\_Light', 'Sports\_Car', 'Subway', 'Wheel', 'Computer', 'Hair', 'Road', 'Automobile', 'Bottle', 'Spring\_Break', 'Car\_Trunk', 'Crystal', 'Smile', 'Stadium', 'TV', 'Couch', 'Art', 'Swimming', 'Moped', 'Ball', 'Surgeon', 'Headrest', 'Van', 'Mobile\_Phone', 'Train', 'Fisheye', 'Bumper', 'Television', 'Cockpit', 'Pc', 'Cell\_Phone', 'Pants', 'Underwear', 'Pool', 'Hat', 'Toilet', 'Diagram', 'Glass', 'Lifejacket', 'Umbrella', 'Alloy\_Wheel', 'Goggles', 'Water', 'Engine', 'Cup', 'Video\_Camera', 'Network'. The query concludes with FAMILY, LINK, THRESHOLD, and MAXITERNUM clauses.

```
SELECT *
FROM GLM (
  ON (SELECT 1)
  PARTITION BY 1
  INPUTTABLE('hackathon.image_inputs_v')
  OUTPUTTABLE('hackathon.glm_outputtable')
  COLUMNS('distraction','Vehicle','Monitor','Sphere','Bag','Screen','Apparel','Teddy_Bear','Hot_Rod','Doctor','X-Ray','Costume','Gemstone','Operating_Scooter','LCD_Screen','Bus','Bra','Deodorant','Seat_Belt','Torso','Kid','Child','Windshield','Light','Cushion','Inflatable','Human','Car_Mirror','Steering_Head','Bowl','Accessories','Female','Mouse','Gearshift','Train_Station','Car_Wheel','Chair','Surgery','Scarf','Phone','Caravan','Black_Hair','Jeans','Me_Highway','Finger','Hippie','Freeway','Plant','Electronics','Hood','Tourist','Shoe','Hoodie','Baby','Potted_Plant','Hospital','Clothing','Blonde','Pillow_Person','Home_Decor','Machine','Sport','Silhouette','Race_Car','Reading','Motor','Denim','Luggage','Console','Girl','Camera','Ankle','People','Hug','Cap','Sunglasses','Sink','Furniture','Overcoat','Cosmetics','Book','Jaguar_Car','Performer','Asleep','Sun_Hat','Car','Dimples','Haircut','Convertible','Laugh_Suit','Lady','Woman','Backpack','Sitting','Face','Glasses','Vest','Offroad','Selfie','Swinwear','Toy','Laptop','Flying','Jacket','Suitcase','Jeep','Port_Footwear','Text','Belt','Beverage','Sedan','Car_Seat','Bikini','Gym','Mirror','Motor_Scooter','Vespa','Fitness','Suv','Arm','Arena','Newborn','Antique_C Traffic_Light','Sports_Car','Subway','Wheel','Computer','Hair','Road','Automobile','Bottle','Spring_Break','Car_Trunk','Crystal','Smile','Stadium','TV','Couch','Art','Swimming','Moped','Ball','Surgeon','Headrest','Van','Mobile_Phone','Train','Fisheye','Bumper','Television','Cockpit','Pc','Cell_Phone','Pants','Underwear','Pool','Hat','Toilet','Diagram','Glass','Lifejacket','Umbrella','Alloy_Wheel','Goggles','Water','Engine','Cup','Video_Camera','Network')
  FAMILY('LOGISTIC')
  LINK('CANONICAL')
  THRESHOLD('0.01')
  MAXITERNUM('10')
);
```

# Aster Analytics



The screenshot shows the Teradata Studio interface. The main window is titled 'Teradata Result Set Viewer' and displays a table of statistical results. The table has 7 columns: an index, 'predictor', 'estimate', 'std\_error', 'z\_score', 'p\_value', and 'significance'. The results include coefficients for 'sphere' and 'bag', model fit statistics like AIC and BIC, and a Wald Test result.

	predictor	estimate	std_error	z_score	p_value	significance
1	(Intercept)	2.11111	0.0277326	76.1235	0.0	***
2	sphere	0.00459449	0.00537346	0.855033	0.392533	
3	bag	-0.00580919	0.0137391	-0.422821	0.672426	
4	screen	-2.74811E-4	0.00298408	-0.0920924	0.926625	
5	ITERATIONS #	4.0	0.0	0.0	0.0	Number of Fisher ...
6	ROWS #	13813.0	0.0	0.0	0.0	Number of rows
7	Residual deviance	9448.16	0.0	0.0	0.0	on 13809 degree...
8	Pearson goodness...	13813.3	0.0	0.0	0.0	on 13809 degree...
9	AIC	9456.16	0.0	0.0	0.0	Akaike informatio...
10	BIC	9486.29	0.0	0.0	0.0	Bayesian informat...
11	Wald Test	5932.04	0.0	0.0	0.0	***
12	Dispersion param...	1.0	0.0	0.0	0.0	Taken to be 1 for ...



# Credits:

TERADATA



Team TERALEXA

# Credits:

TERADATA



Team TERALEXA

# Credits:

## Team XPRESSION:

Graham

JSON

Melissa

Luyanda

Andrew

Karan

Tam

## Images and Presentation Templates:

[mlive.com](http://mlive.com)

[CarnivalSlides.com](http://CarnivalSlides.com)

[BusinessInsider.com](http://BusinessInsider.com)

A wide-angle photograph of a long, straight asphalt road that stretches from the foreground into the distance, vanishing at the horizon. The road is flanked by rolling hills covered in sparse, dry vegetation and patches of snow. The sky is overcast with heavy, grey clouds. The overall mood is desolate and expansive.

// .XPRESSION



A photograph of a long, straight asphalt road stretching into the distance. The road is flanked by hills with patches of snow. The sky is overcast and grey. The text "The Road to Come" is overlaid in the bottom right corner in a teal color.

The Road to Come