MOMMA
Modeling of Measurable Maternal Attributes

Diana Amador

Andrew Dempsey

Samantha Godfrey

Peter Steinhoff

MATH 579: Mathematical Modeling
Dr. Jen-Mei Chang
May 12, 2011

Research Topics

- Study possible connections between maternal conditions and placenta weight or placental ratio
- Implementation of Statistical Analysis, Principal Component Analysis, and Logistic Regression
- Study different indicators of placental growth

Placenta Weight

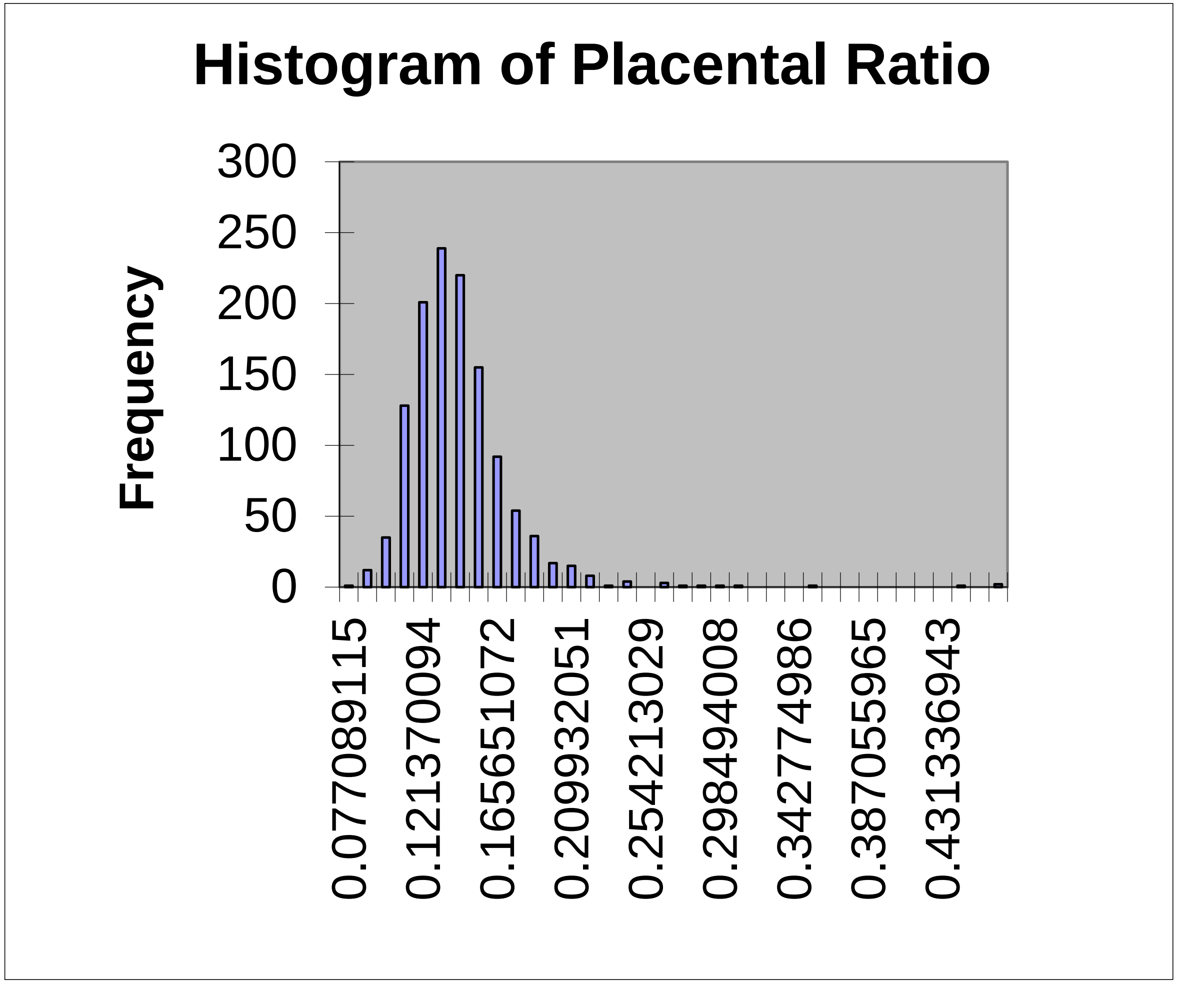
- Placenta weight is a factor that reflects the placental development
- The weight of the placenta can be an indicator of fetal abnormalities
- We analyzed the maternal characteristics that were provided to see if they had any significant effect on placental weight

Maternal Variables

- Maternal Pre-Pregnancy BMI
- Maternal Age
- Mother’s Ethnicity
- Gestational Diabetes

What is Placental Ratio?

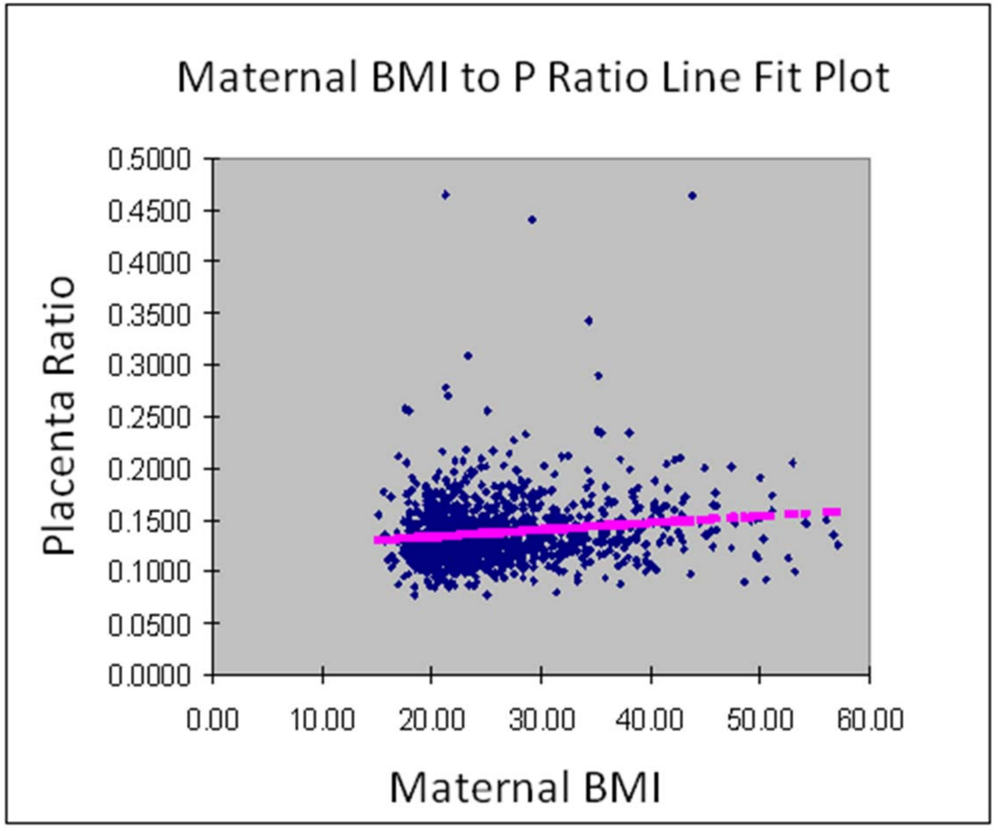
$$\text{Placental Ratio} = \frac{\text{Placental Weight}}{\text{Birth Weight}}$$



Statistical Analysis

- We conducted separate studies for each maternal variable
  - Maternal Pre-Pregnancy BMI
  - Maternal Age
  - Mother’s Ethnicity
  - Gestational Diabetes
- We compared the statistical results to previous studies

Maternal Pre-Pregnancy BMI and Placenta Ratio



Statistical Results

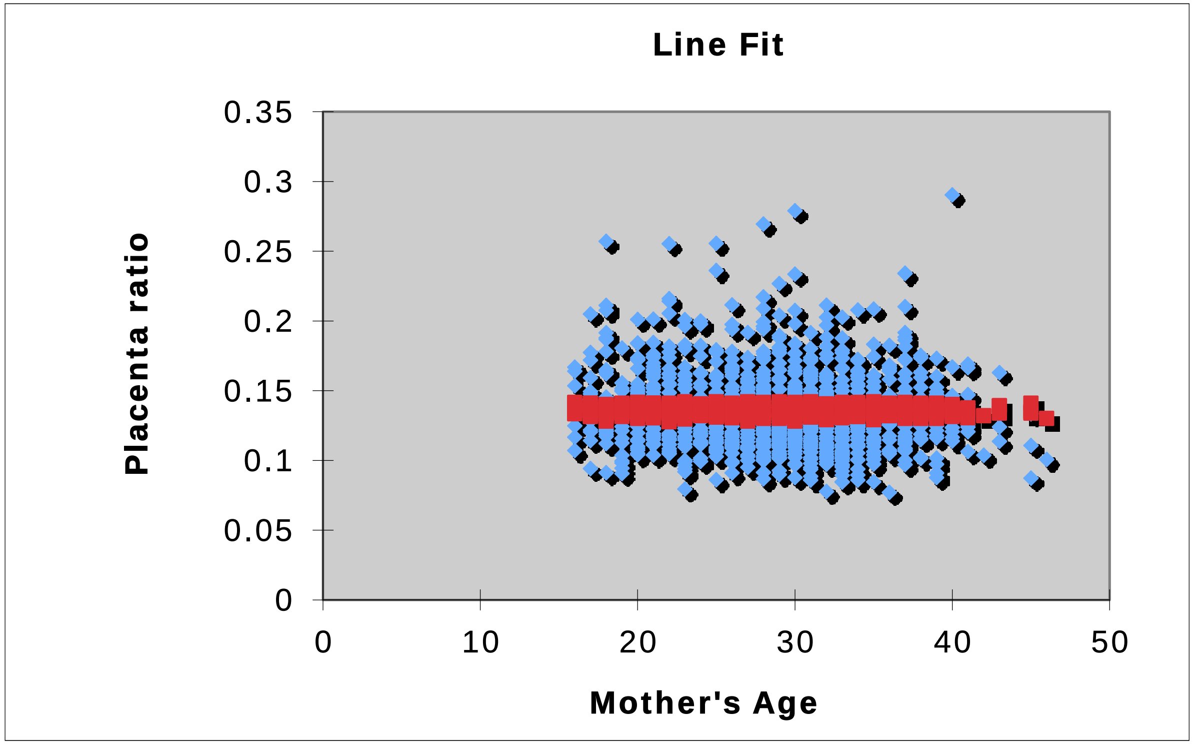
R^2 value: 0.0209

F-stat P-value: 3.54E-07

Conclusions

- Although a positive linear correlation exists between BMI and the Placenta Ratio, the relationship is too weak to use as a predictor of Placenta Ratio and hence Placenta weight or baby weight.

Maternal Age and Placenta Ratio



Statistical Results

R^2 value: 0.01

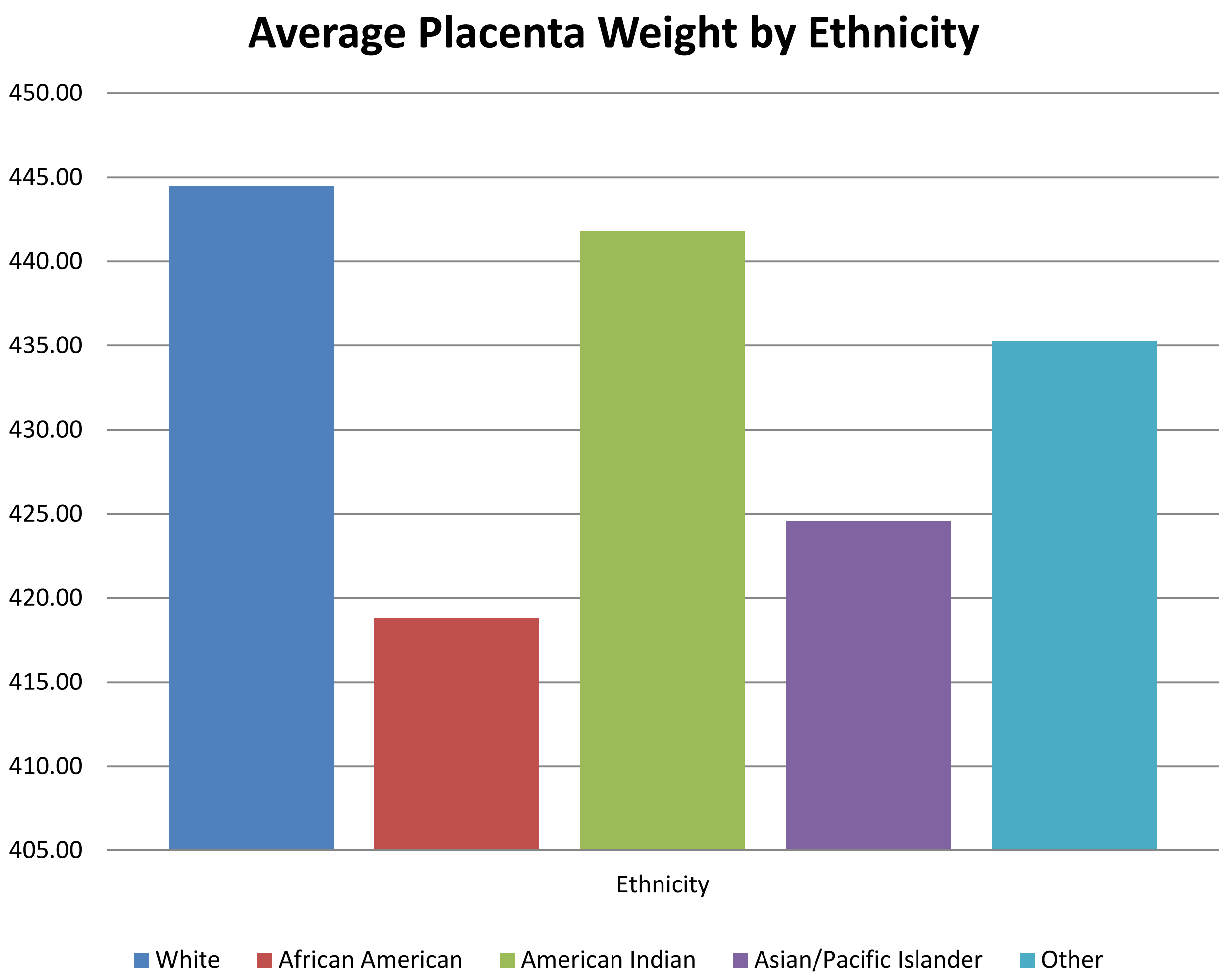
F-stat P-value: 0.0005

Conclusions

• There is a statistically significant correlation between maternal age and placenta ratio. However, the relationship is so weak that a linear fit is basically useless.

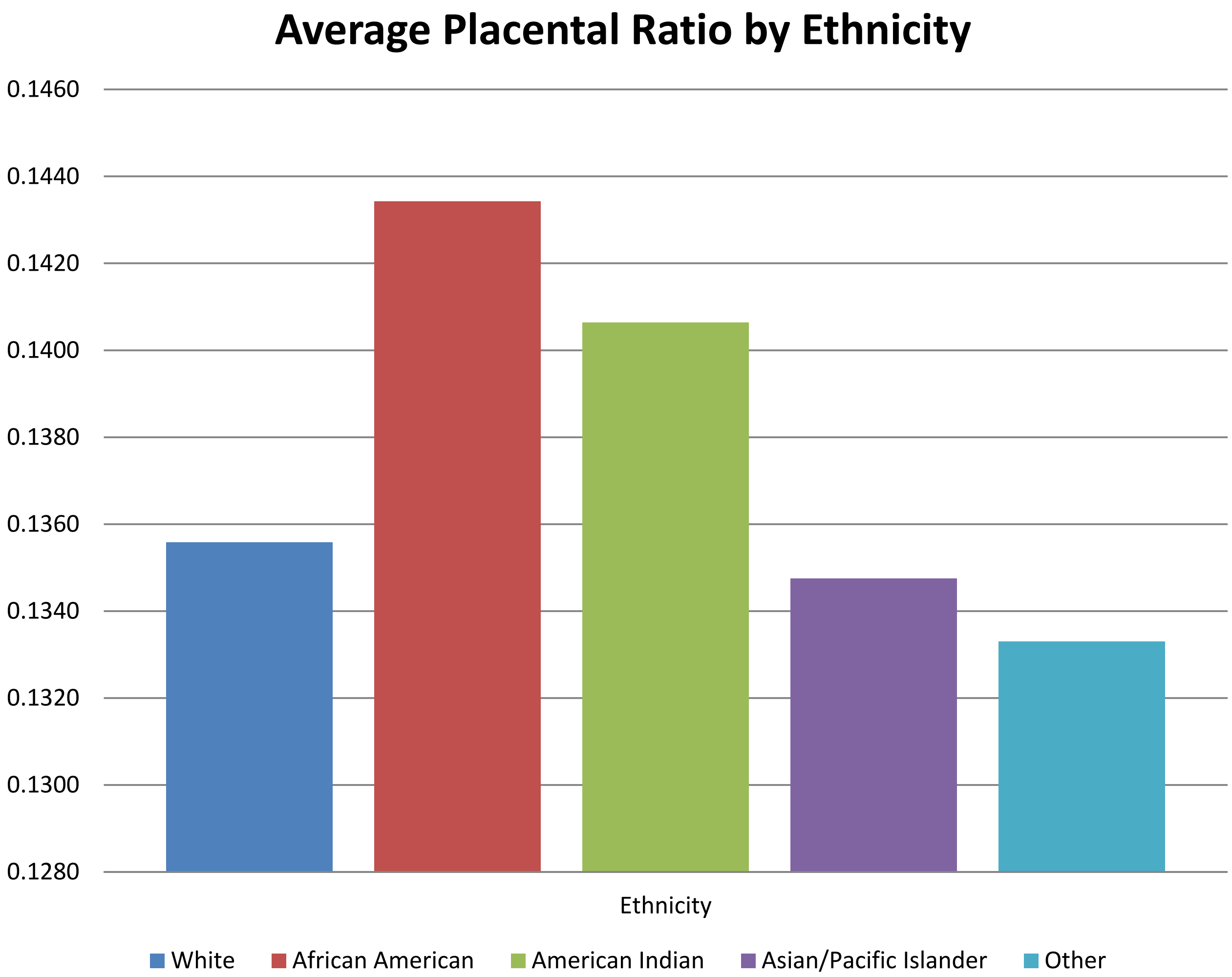
Mother’s Ethnicity

- Highest average placenta weight with white mothers (449.49±101.94)
- Lowest average placenta weight with African American mothers (418.81±115.79)
- African Americans having the lowest placental weight is supported by past studies [1][2]

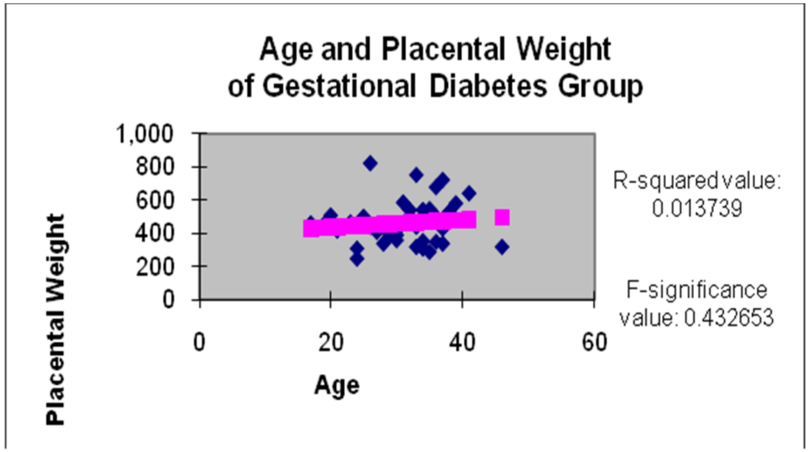


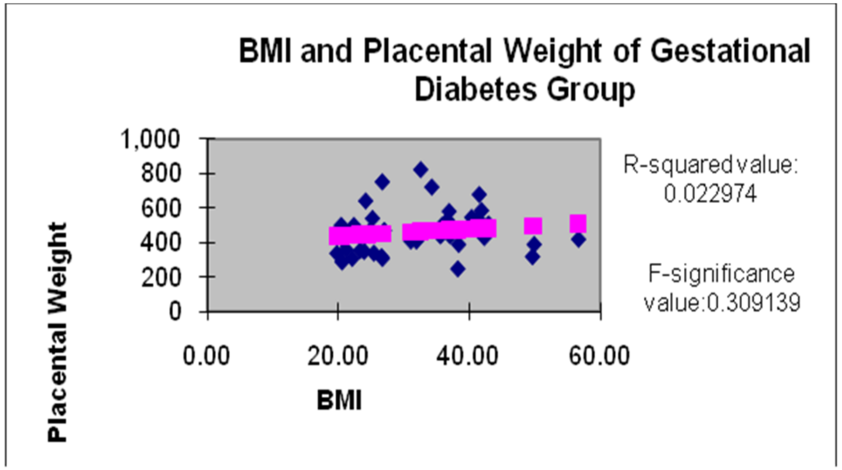
Mother’s Ethnicity

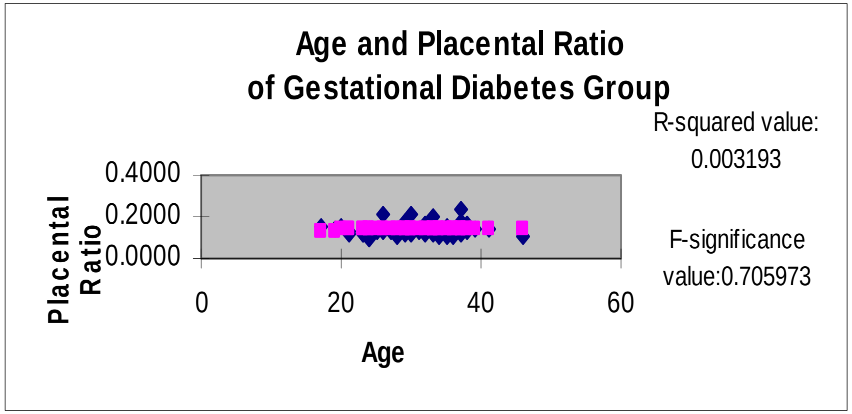
- Highest placental ratio with African American mothers (.1434±.0331)
- African American babies in this study had the lowest average birth weight compared to their placenta weight
- Average placental ratio ranged from 13.3% to 14.4%

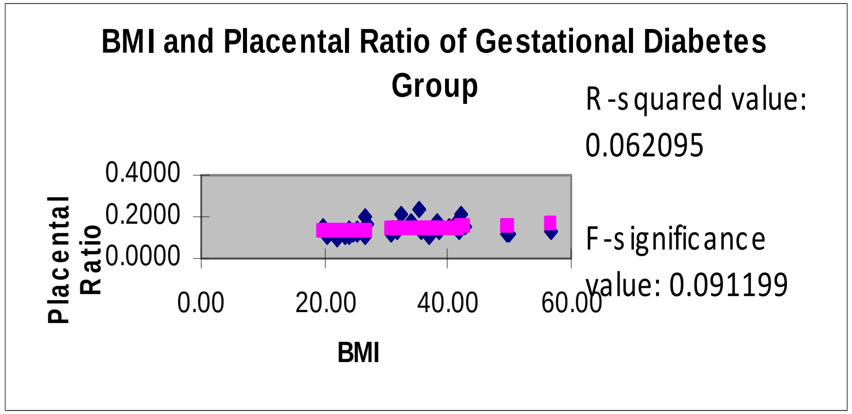


Gestational Diabetes









Conclusion: Placental Weight/Ratio is independent of gestational diabetes

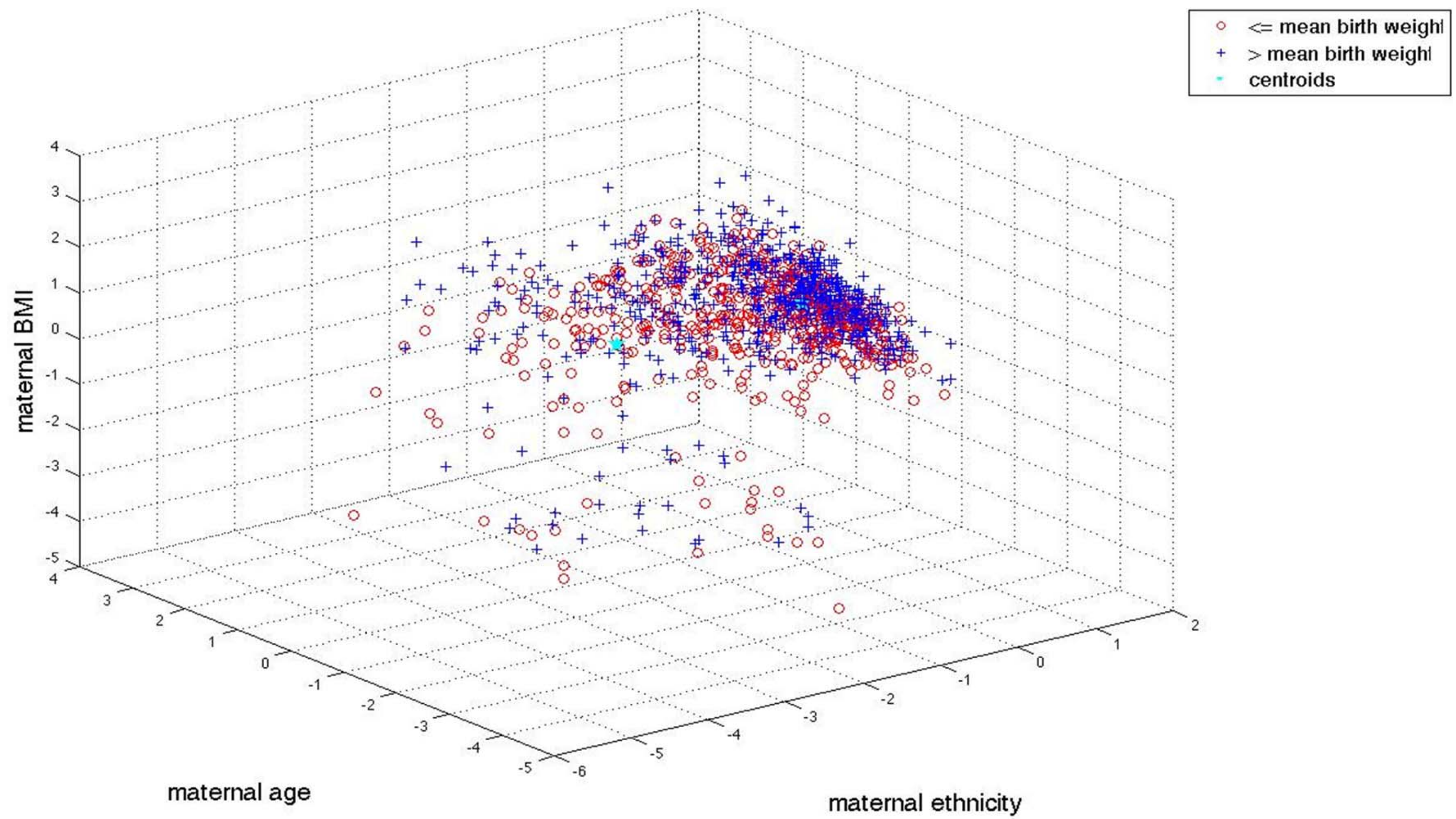
K-means Clustering

- K-means is an iterative method that attempts to separate a set of *n* observations into *k* clusters

- We were looking for clustering within various *characteristic spaces* driven by a specified "ground-truth" maternal trait

- Well defined clustering would indicate a highly influential ground-truth trait.

Birth weight clustering?



Goodness-of-Fit

- Q: Assuming *some* clustering occurs, how can we quantify it?
  - We assigned identities to the clusters based on the ground-truth of their populations
  - We then compared each observation’s ground-truth to that of the closest cluster centroid, and counted the number of positive matches.
  - A highly homogeneous cluster would have a large number of positive matches.

Mahalanobis Distance

- An observation's "closest" cluster may depend upon the distance metric used.

- The Mahalanobis metric compensates for differences in axes scaling
  - Distortion arises from our use of different characteristics to define the vector space

$$D_m(x) = \sqrt{(x-\mu)^T S^{-1} (x-\mu)}$$

- *x* is the observation,µ is the cluster centroid, and *S* is the cluster's covariant matrix.

Principal Component Analysis

What's it good for?

- Multivariable Data Sets
- Variable Reduction
- Visual "Representation" of N-dimensions
- Projection down onto 3D or 2D spaces

Basics of PCA

1) Normalize Data Set
2) Find the Covariance Matrix of the Normalized Matrix
3) Find eigenvalues and eigenvectors of the Covariance Matrix
4) Locate Principal Components by finding the eigenvectors with the highest eigenvalues
5) Form a "Feature" matrix by using the chosen eigenvectors as columns
6) Transpose the feature matrix into FM and transpose the normalized matrix into NM
7) Result Matrix = FM x NM
8) Hope you will be able to interpret the results…

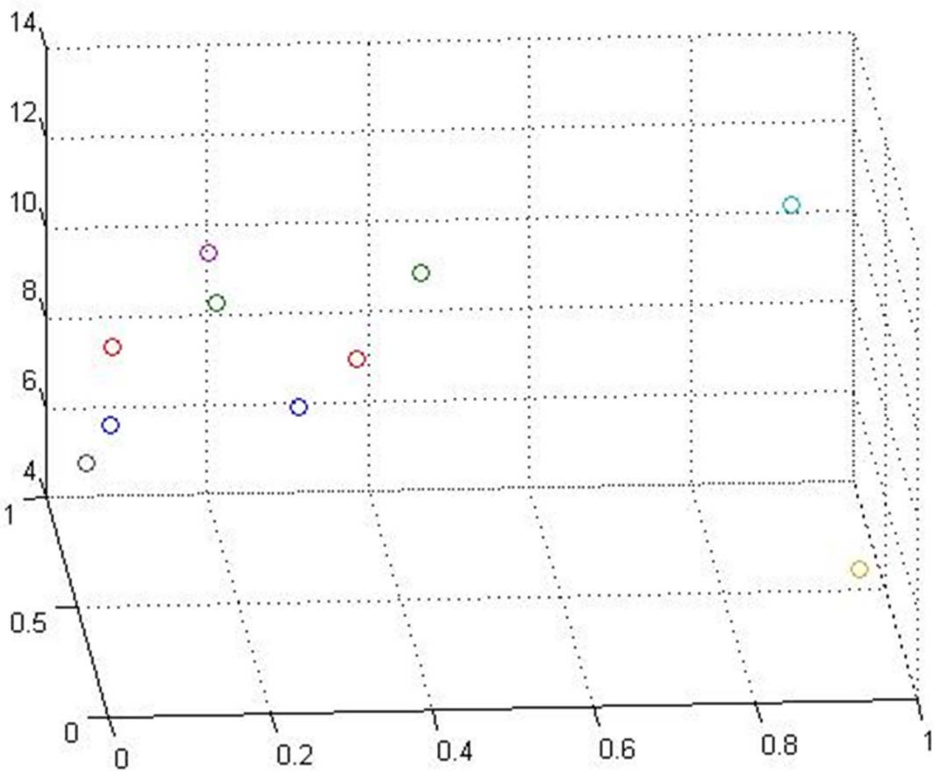
Coordinated Data VS. Uncoordinated Data

X	Y	Z1
0.05875	0.7127	7.00000
0.39874	0.1518	13.00000
0.89338	0.6273	12.00000
0.02306	0.2498	11.00000
0.12820	0.0547	14.00000
0.94468	0.1850	6.00000
0.04631	0.9467	5.00000
0.31202	0.9809	6.00000
0.18357	0.6406	10.00000
0.38509	0.9883	7.00000

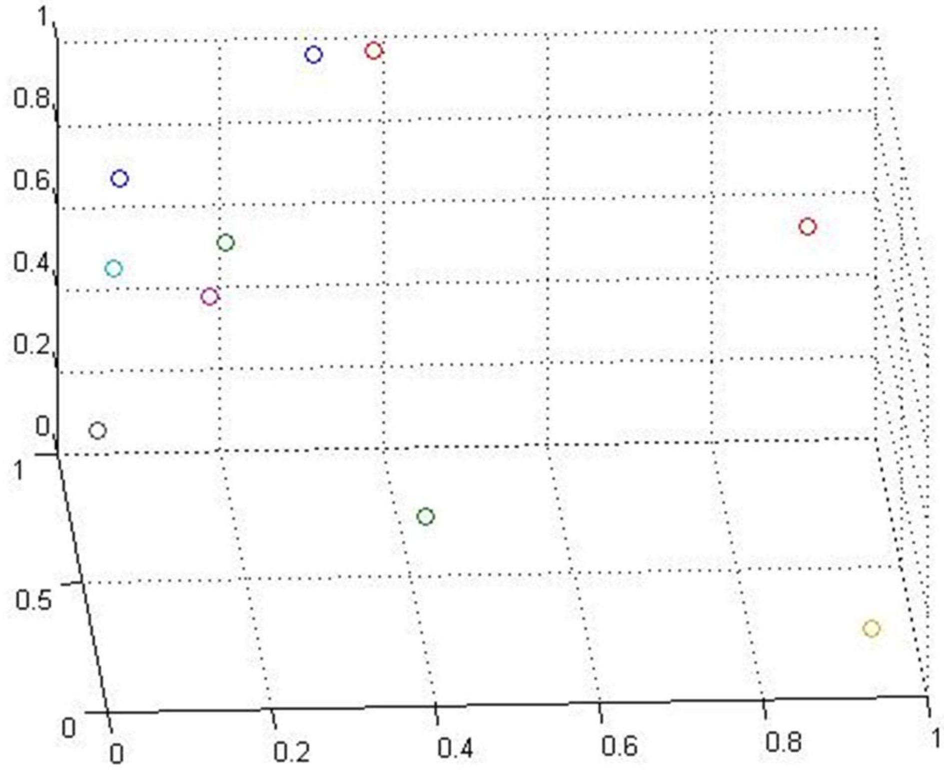
VS.

X	Y	Z2
0.05875	0.7127	0.84778
0.39874	0.1518	0.36390
0.89338	0.6273	0.75461
0.02306	0.2498	0.91804
0.12820	0.0547	0.96854
0.94468	0.1850	0.05021
0.04631	0.9467	0.09188
0.31202	0.9809	0.97453
0.18357	0.6406	0.73535
0.38509	0.9883	0.97415

3D plot of previous slide



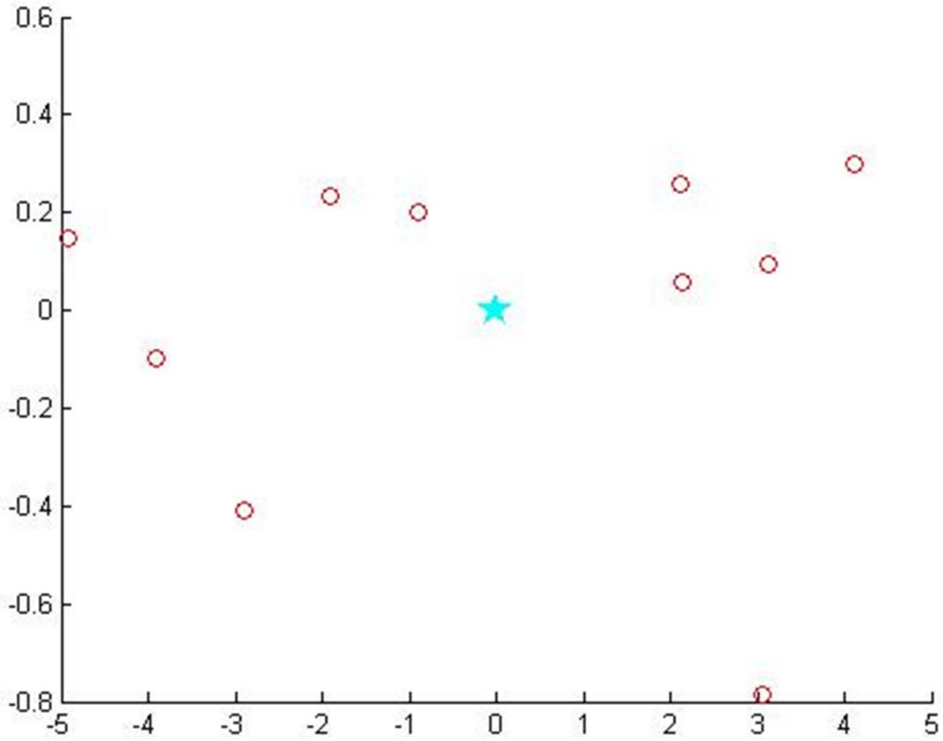
(X,Y,Z1)



(X,Y,Z2)

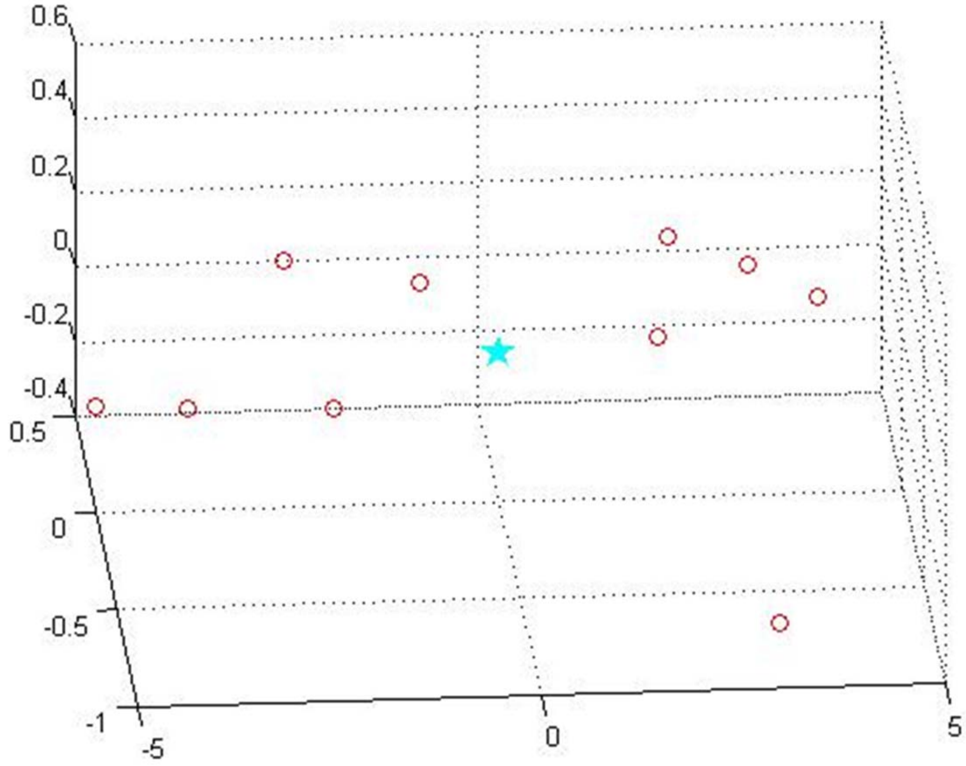
Coordinated Data Result

2D



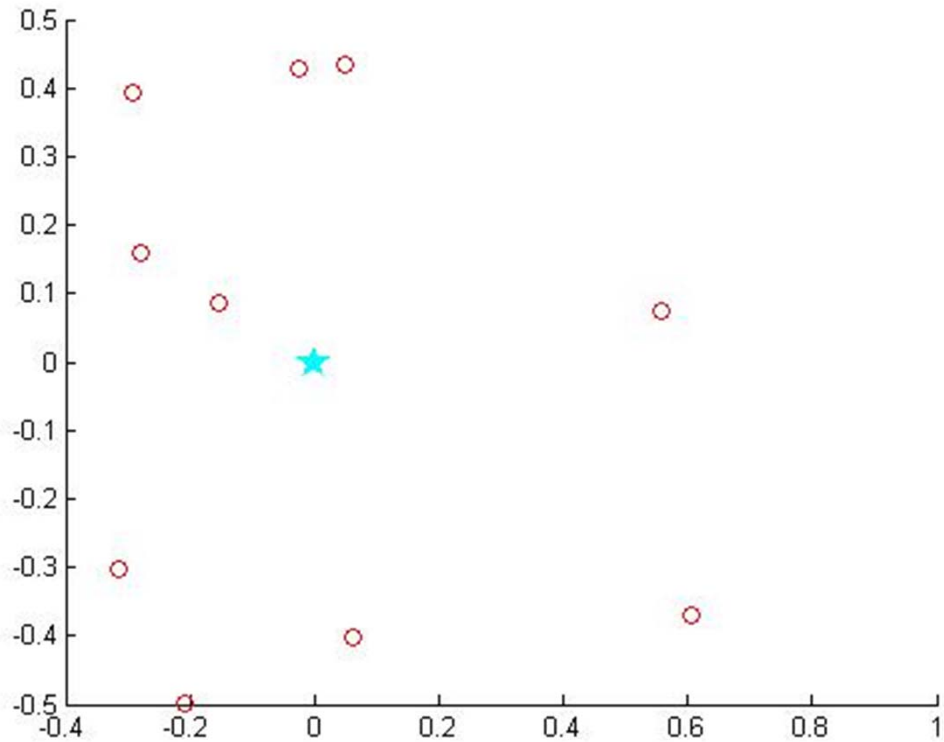
Coordinated Data Result

3D



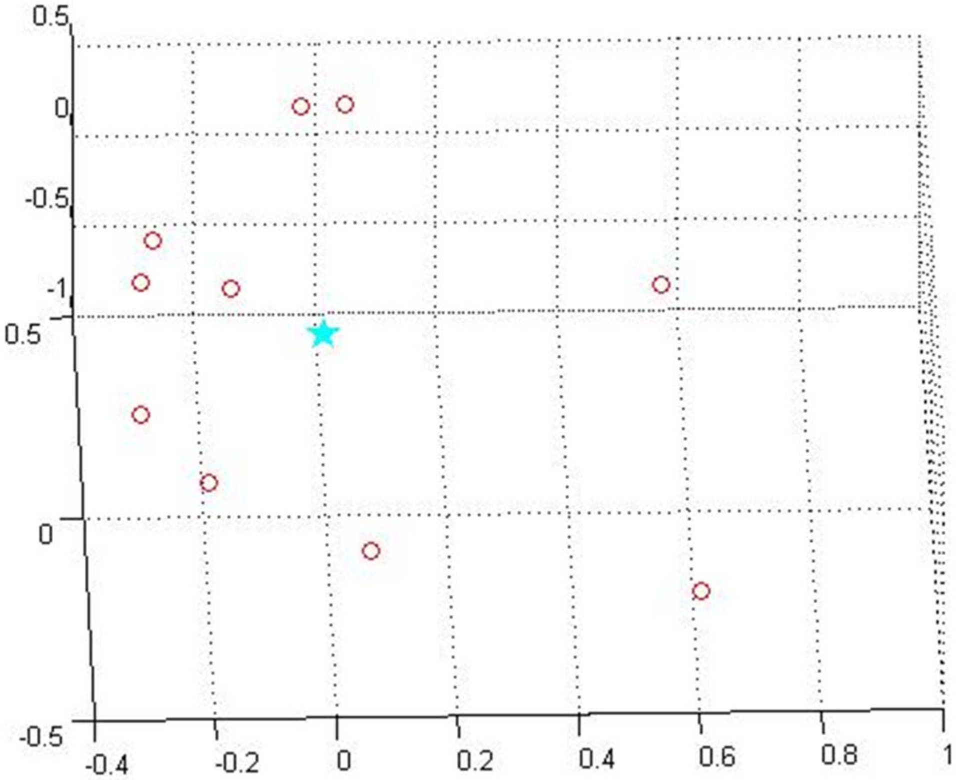
Uncoordinated Data Result

2D



Uncoordinated Data Result

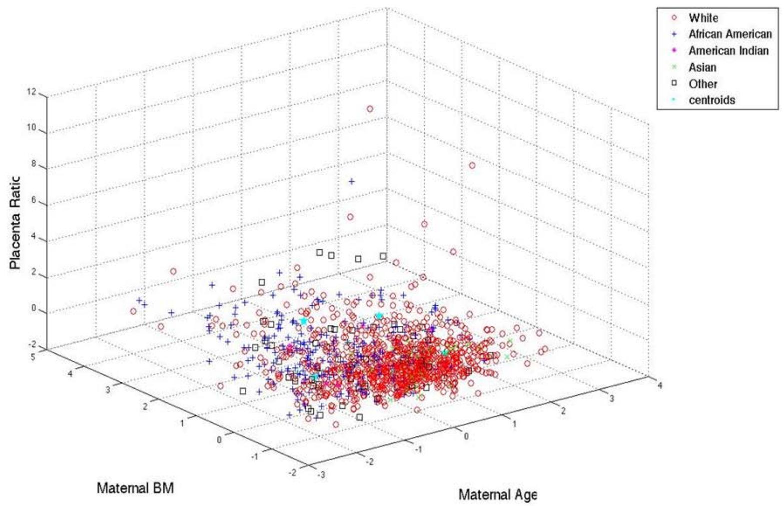
3D



K-means clustering *with* PCA

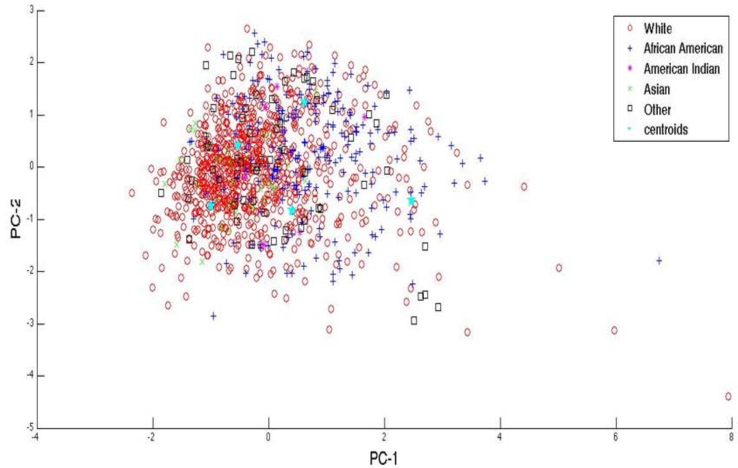
- PCA projects the observation vectors into a principal component subspace, *typically* with a reduction in dimension.
- In the case of higher dimensional characteristic spaces, PCA can project the data into R^3, R^2, or R.
- By treating the data with PCA prior to clustering we hope to add clarity to the whatever separation might be present.
- The ground-truth coloring stays the same.

Ethnicity Clustering?



A 3-D characteristic space projected into a 2-D principal component subspace

The ground-truth coloring is the five ethnicities defined by the maternal data



Select Results

case	***Ethnicity 1.5***	***Age 1.0***	***Birth weight 1.5***
spatial dimensions *n*	2	5	5
spatial characteristics	maternal BMI, placenta ratio	maternal ethnicity, maternal BMI, gestational diabetes, placenta weight, birth weight	maternal ethnicity, maternal age, maternal BMI, gestational diabetes, placenta weight
ground-truth	ethnicity	maternal age	Birth weight
number of clusters *k*	5	7	2
GoF under random assignment	20%	17%	50%
GoF k-means clustering in R^n (Euclidean)	9%	28%	55%
GoF k-means clustering in R^n (Mahalanobis)	9%	27%	55%
Principal Components	1	3	2
GoF k-means in PC-subspace (Euclidean)	31%	22%	48%
GoF k-means in PC-subspace (Mahalanobis)	31%	21%	48%

Logistic Regression

- Uses a binomial distributed set of data to predict a success/fail outcome based on several variable combinations
- Useful way of describing the relationship between one or more independent variables and a binary response variable expressed as a probability that has only two values.
- Determine the probability that a certain event will happen with different combinations of independent risk variables.

Logistic Regression

The odds ratio, R, is the ratio of the probability that a certain event will happen to the probability that it will not happen. [3]

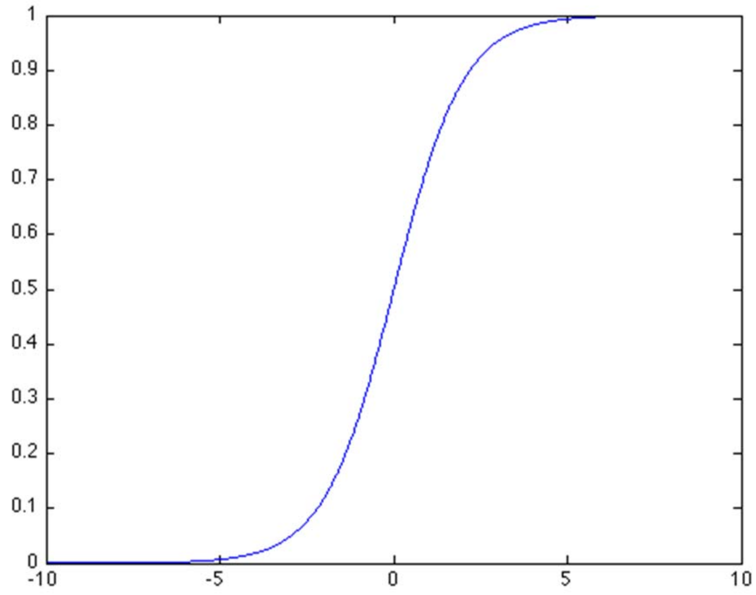
Let $R = e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p}$

$$R = \frac{P}{1-P}$$

$$P = \frac{R}{1+R}$$

Let $z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$

Then $P = \dfrac{1}{1+e^{-z}}$



Logistic Regression

- Goal: Find logistic regression coefficients

  β_0 , β_1 , β_2 ,…,β_p

- Matlab built in function glmfit

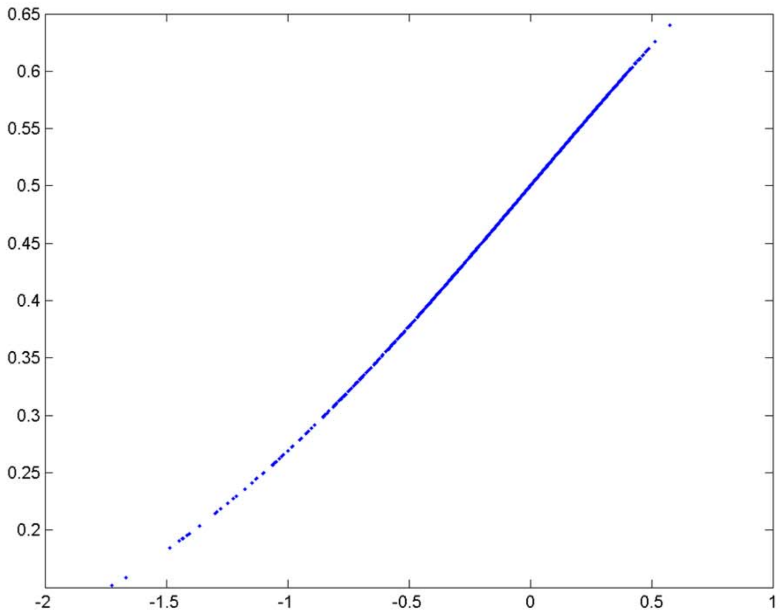
  B=glmfit(X, Y, ’binomial’, ’link’, ’logit’)

  P=glmval(B, x, ‘logit’)

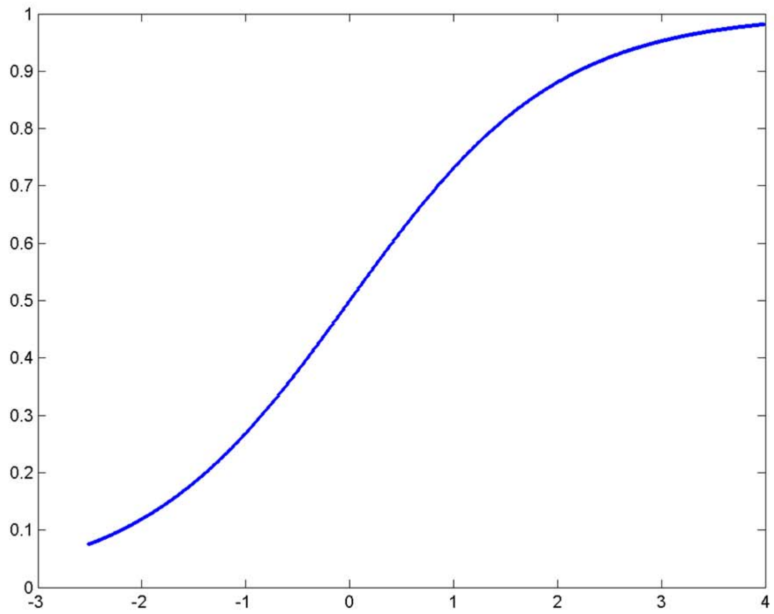
Logistic Regression: β-value

- "Success" situation: β<.75 since the placenta is considered more efficient for these values

  β_0 =-.739364, β_1 =.050368, β_2 =-.017877, β_3 =-.066001, β_4 =.048526



Plot from our data set

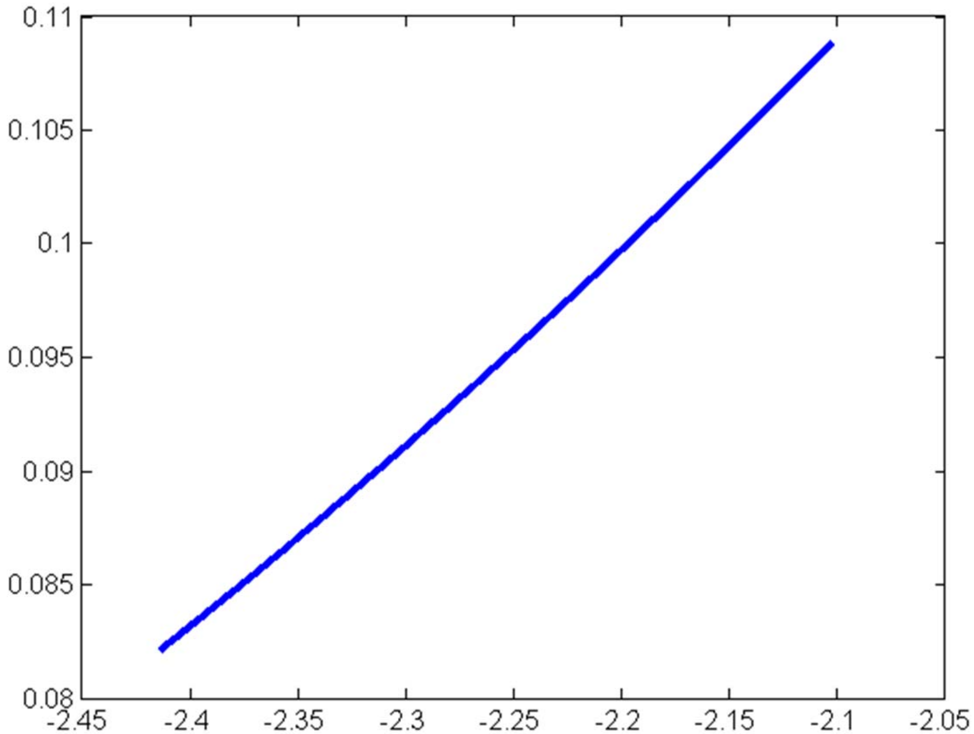


Extended plot for values outside our data set

Logistic Regression: Low Birth Weight

- "Success" situation: low birth weight (<2500 g)

  β_0 =-2.118044, β_1 =.0150295, β_2 =-.0196460, β_3 =.3892553, β_4 =.0120513



Extended plot for values outside our data set

Conclusions & Further Work

- Across our various methods, results were often statistically significant, but too weak to be used for prediction.
- Little to no estimation of placenta weight can be deduced from the following characteristics: maternal ethnicity, age, BMI, or diabetic status.
- A covariance matrix between maternal characteristics and principal components might pinpoint the most influential principal components.
- The data under consideration could be expanded to include additional maternal and placental characteristics.

References

[1] Greg R. Alexander, Michael Kogan, Deren Bader, Wally Carlo, Marilee Aleen, and Joanne More. US birth weight/gestational age-specific neonatal mortality: 1995-1997 rates for whites, hispanics, and blacks. *Pediatrics,* 111(1):61-66, 2003.

[2] Kesha Baptiste-Roberts, Carolyn M. Salafia, Wanda K. Nicholson, Anne Duggan, Nae_Yuh Wang, and Frederick L. Brancati. Maternal risk factors for abnormal placental growth: The national collaborative perinatal project. *British Medical Journal Pregnancy and Childbirth,* 8(44), 2008.

[3] James R. Frederick. Logistic Regression. http://www.uncp.edu/home/frederick/DSC510/LogisticReg.htm, March 2003.

[4] C.M. Salafia and M. Yampolsky. Metabolic scaling law for fetus and placenta. *Placenta,* 30:468-471, 2009.