Autism Risk Behaviors Study

Rose Din, Lisa Kaunitz, Claire Markey, Allison Ocelo, Cristina Sanchez, Kayla Schroeder

Research Questions

1. What are the significant risk factors that predispose a child to ASD behaviors?

2. Is there a significant difference when looking at predictors for autism in males vs females?

Literature Review - Gender Differences

Males

- ASD can be diagnosed in boys as early as 2 years of age
- 3 males have autism for every female
- Most ASD research was done on male patients

Females

- On average, it may take up to 14 years for girls to be diagnosed with ASD
- Young girls tend to achieve more basic milestones
- Overall better at social skills & peer relationships
- Disrupts the typical bias and norms set in the standards for ASD

Variables of the Study

Prenatal Variables

- Moms delivery age
- Moms education
- Sex
- Art and Art Methods (1-5)
- Mom trauma
- Maternal pregnancy problems (1-21)
- Maternal mental health (1-7) and conditions
- Mom fever & When fever (1-3)
- Drug variables (tobacco, cannabis, alcohol, etc)
- Delivery
- “One” count variables

New Variables

- Maternal health count = count of maternal mental health variables (1-7)
- Tri tobacco
- Tri alcohol
- Tri vaping
- Tri cannabis
- Tri stimulants
- Tri opioids
- Tri psychoactives

Total number of trimesters drug was used

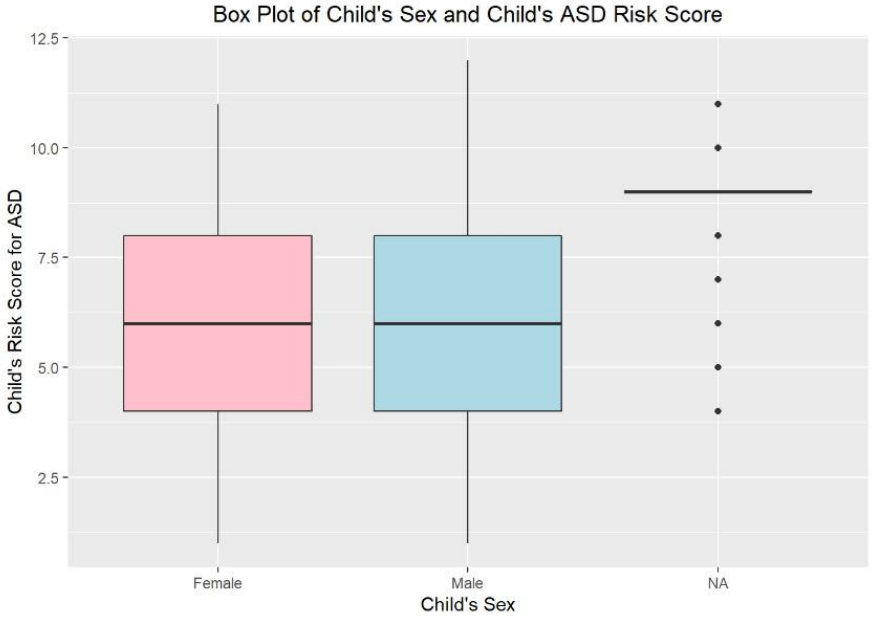
- Baby mortality = difference of “totalpregnacies” and “momslivebirths.”

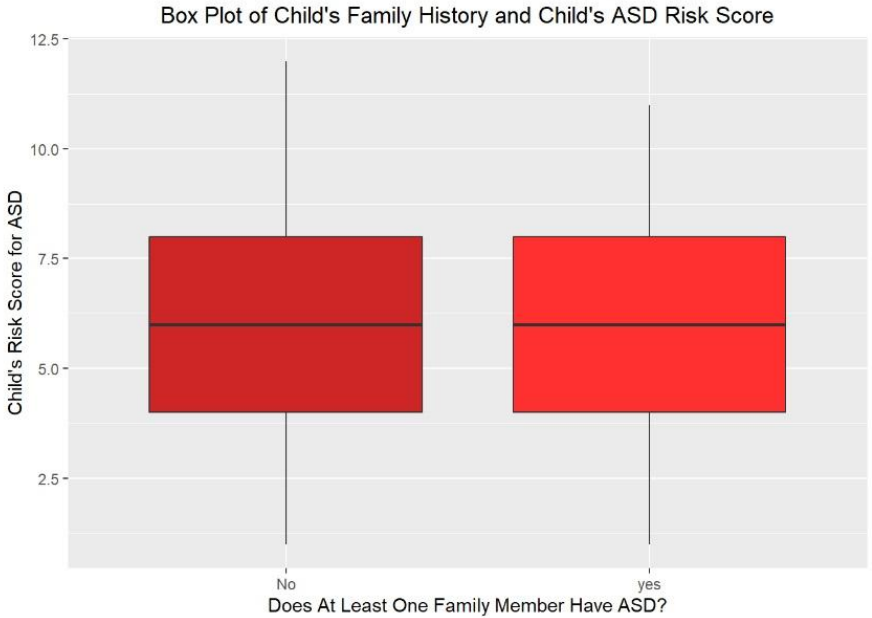
Exploratory Data Analysis

5 Risk Factors According to Mayo Clinic

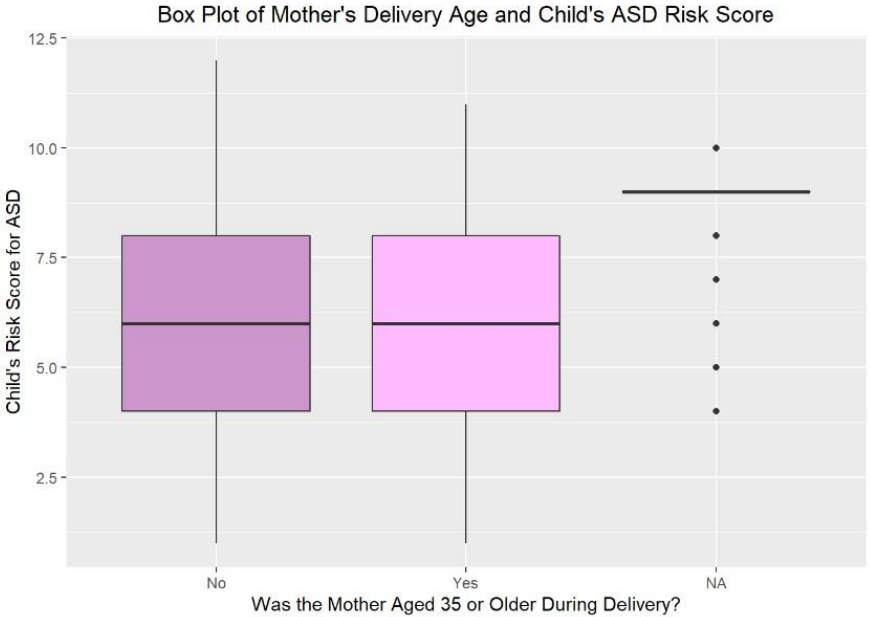
- Child's Sex (boys are more likely than girls to be diagnosed)
- Family History (if anyone related to the baby has already been diagnosed)
- Other Disorders (such as fragile X syndrome, etc.)
- Extremely Preterm Babies (mainly those born before 26 weeks of gestation)
- Parents' Ages (older parents at time of delivery)

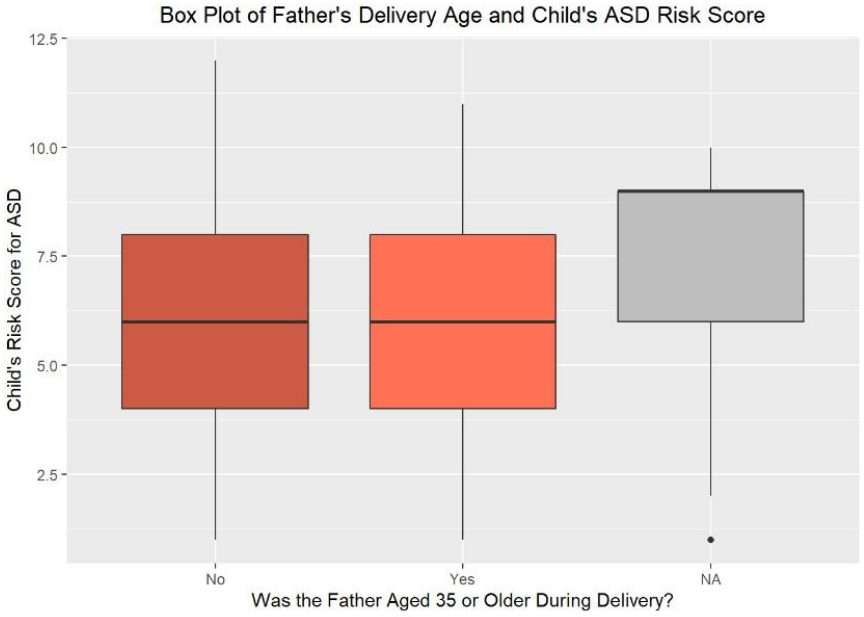
Exploratory Data Analysis cont.



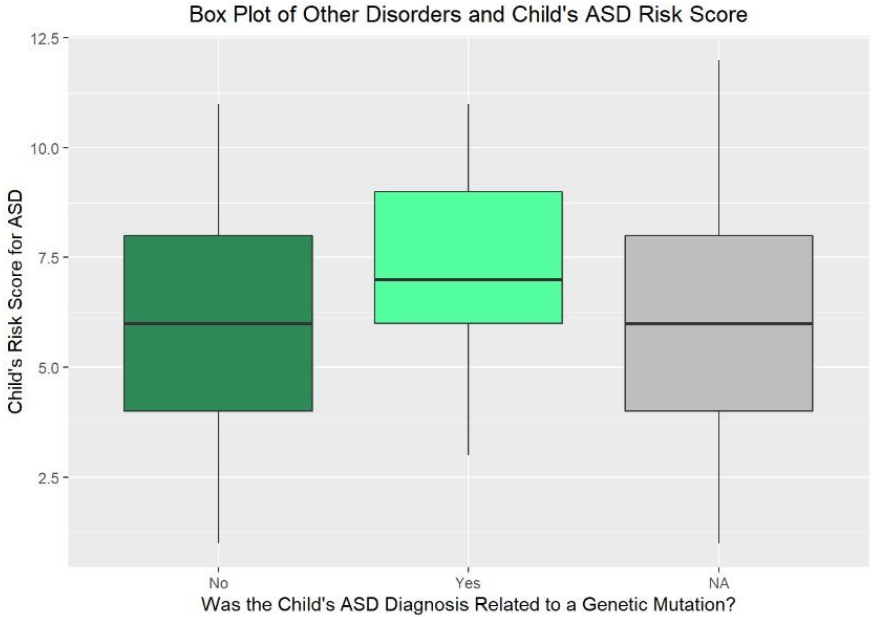


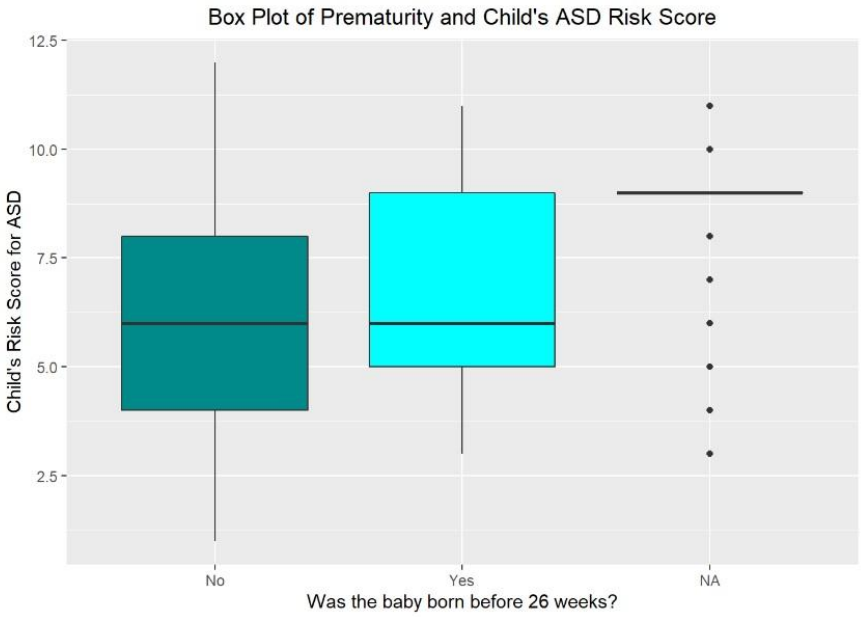
Exploratory Data Analysis cont.





Exploratory Data Analysis cont.

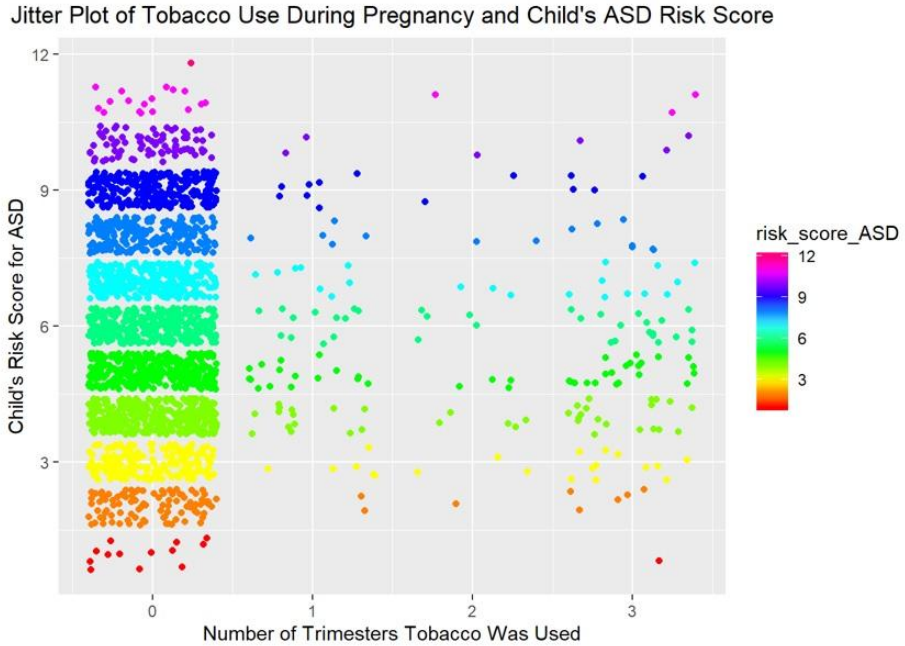




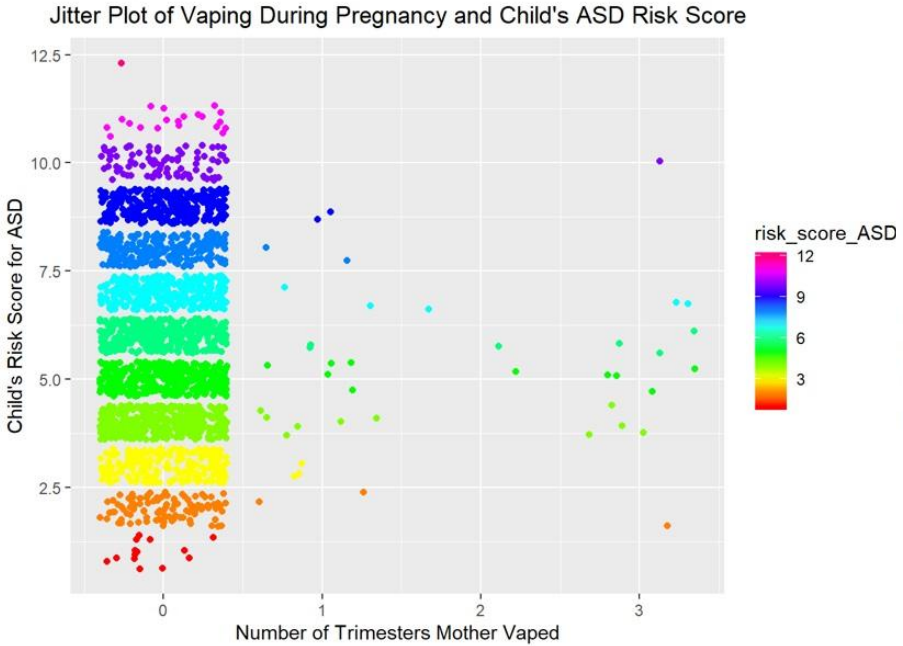
Exploratory Data Analysis cont.

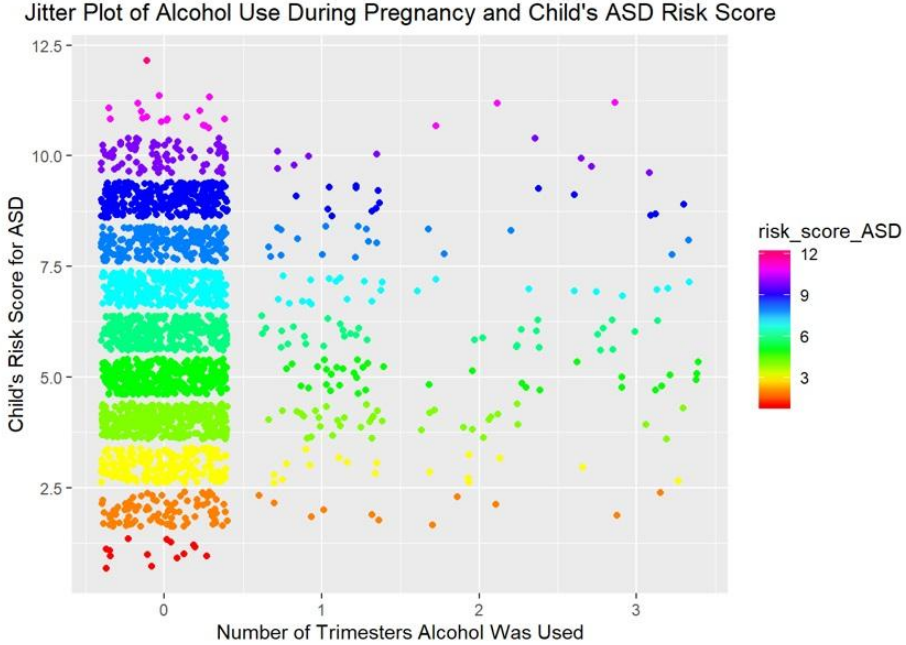
Drug Use During Pregnancy

- Tobacco
- Vaping
- Alcohol
- Cannabis
- Stimulants
- Opioids
- Psychoactives

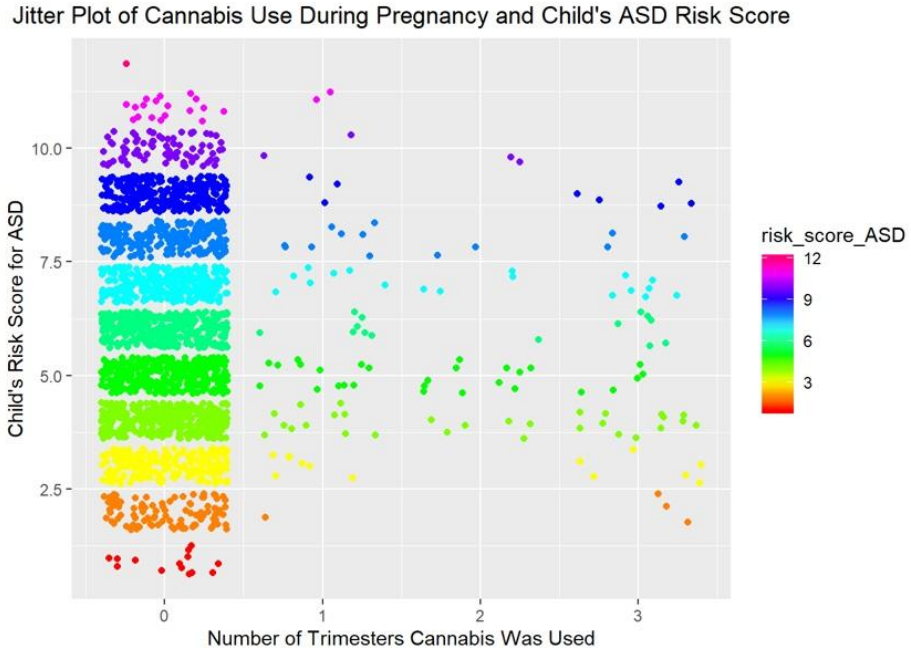


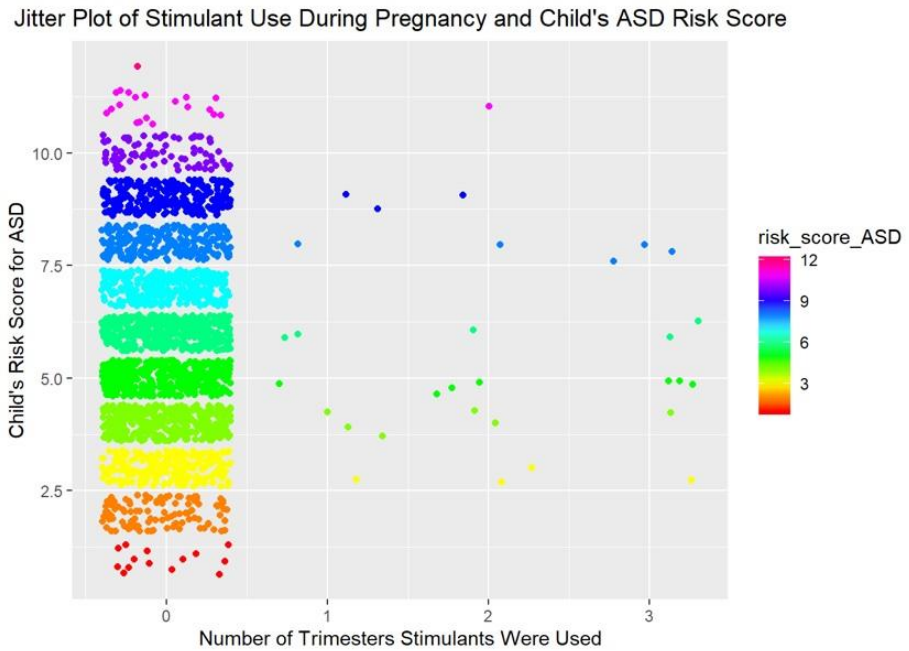
Exploratory Data Analysis cont.



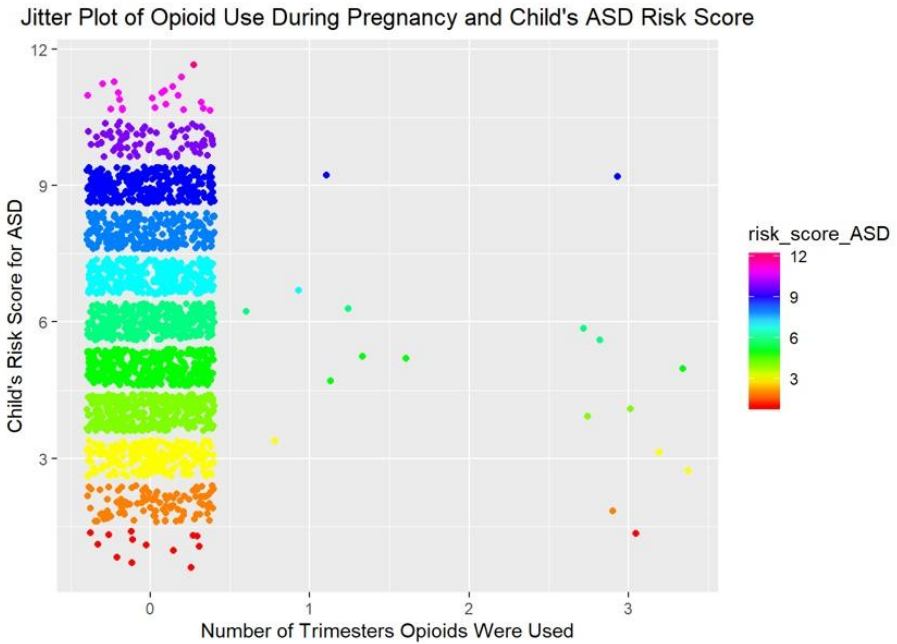


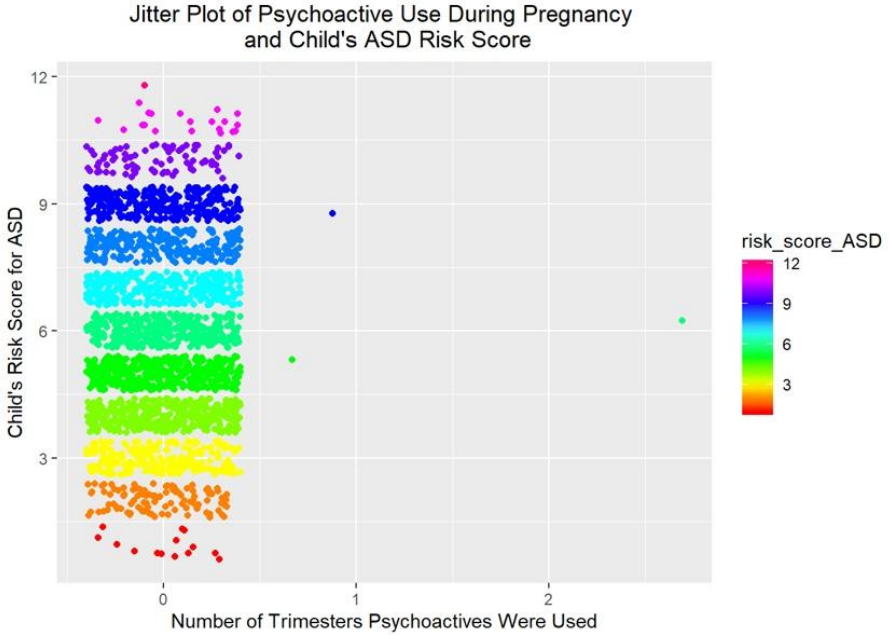
Exploratory Data Analysis cont.





Exploratory Data Analysis cont.

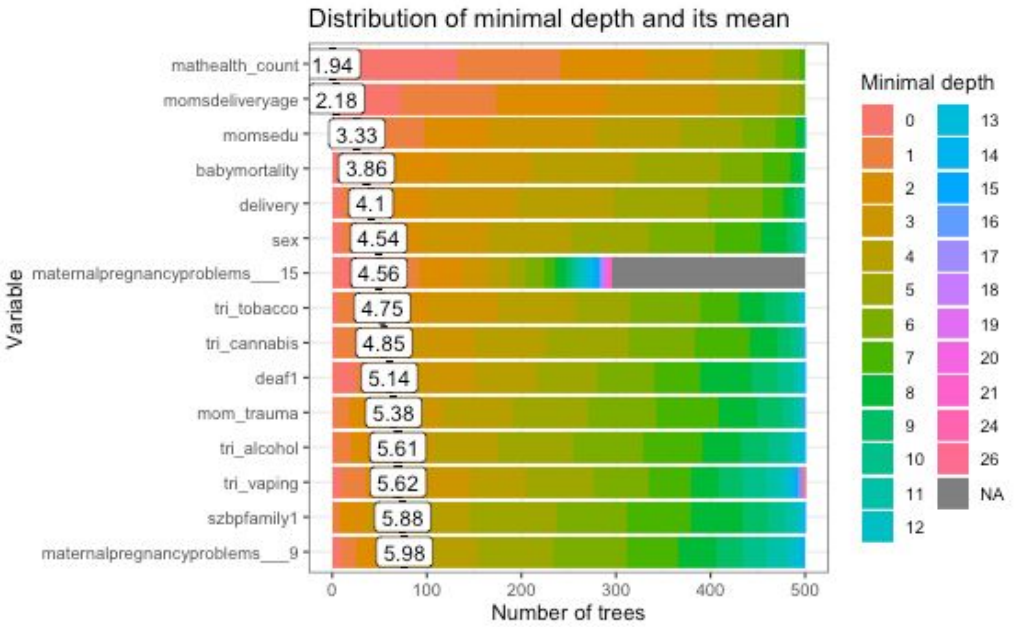




Statistical Analysis

Variable Selection: Random Forest Model

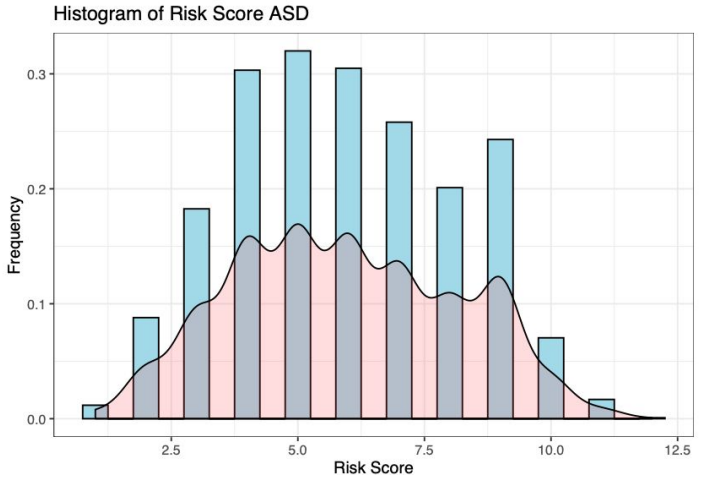
- Node purity and mean squared error measures to determine important variables
- 14 total variables were selected as important risk factors from `risk_score_ASD`



Statistical Analysis

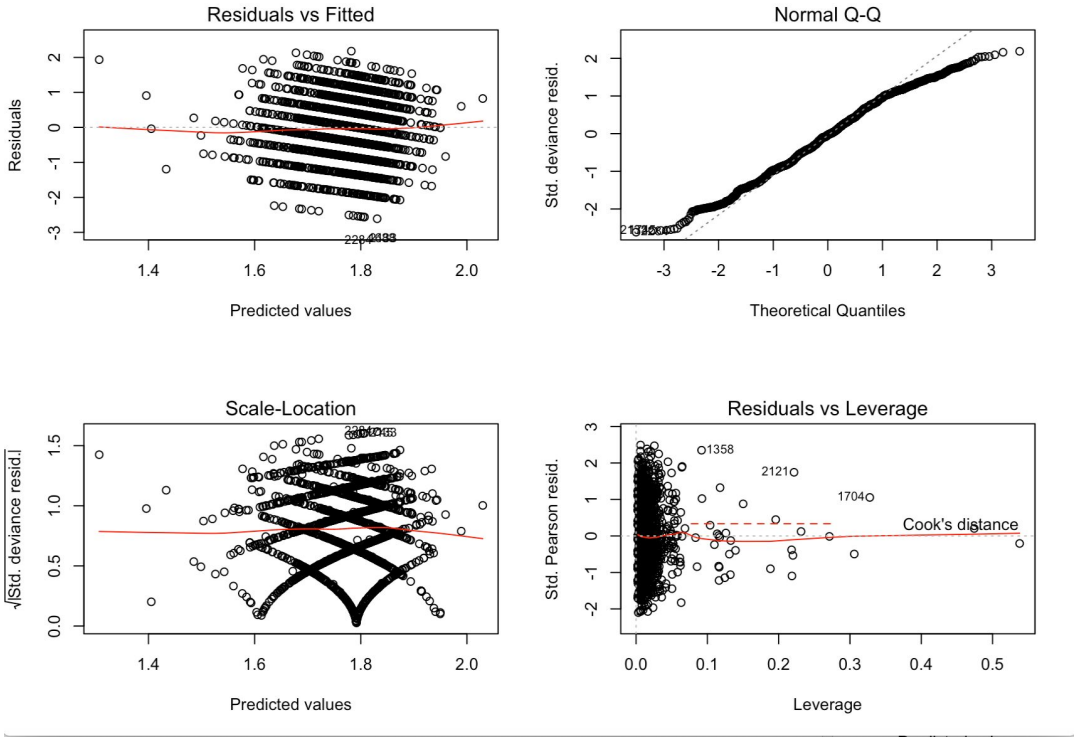
Model Selection: Poisson Regression

- Assumptions
  - Outcome variable consists of count data
  - One or more independent variables
  - Dispersion parameter is less than one
  - Independence of observations
  - Counts follow poisson distribution

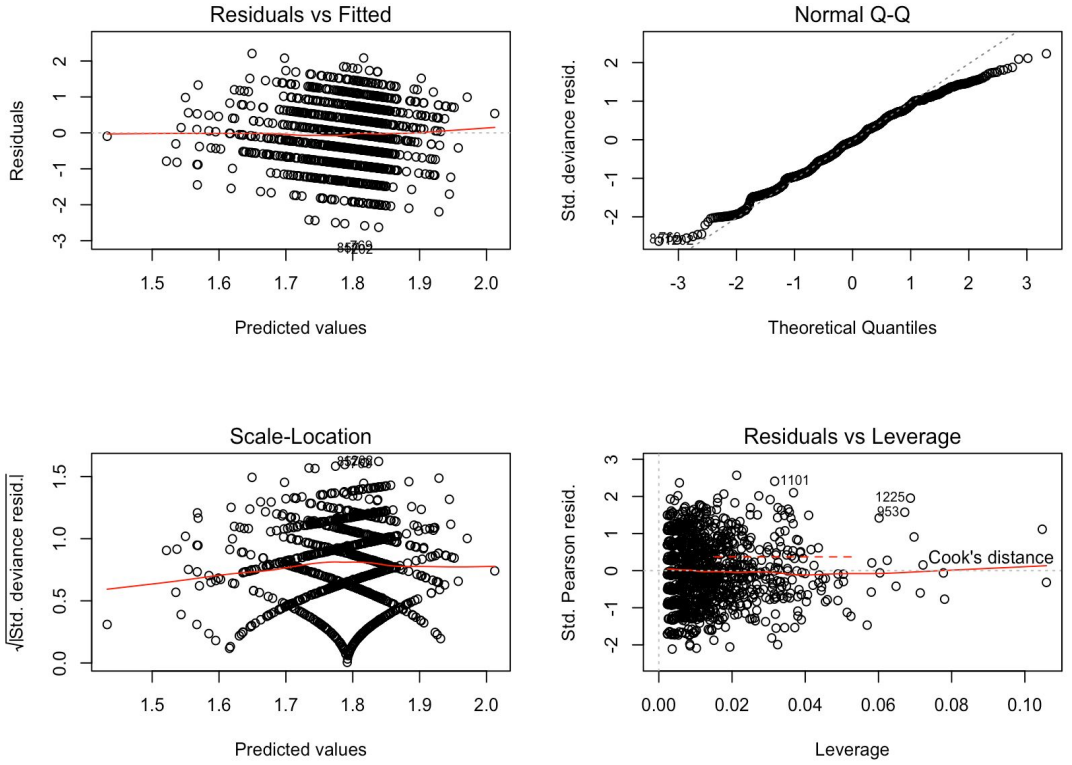


Model	Dispersion Parameters	AIC
Model 1 (full)	0.798	9886
Model 2 (male)	0.748	5167
Model 3 (female)	0.836	4719

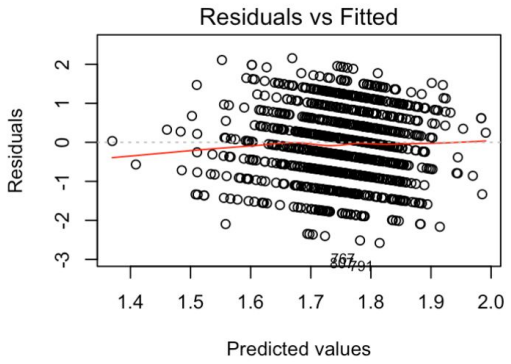
Residual Plots: Full Model

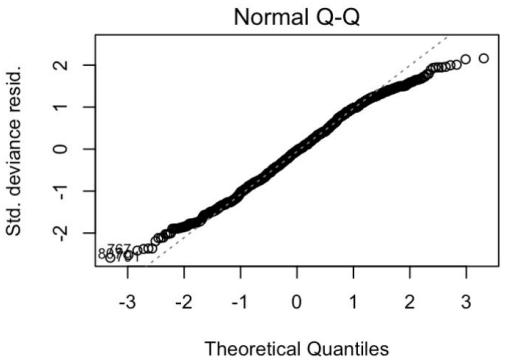


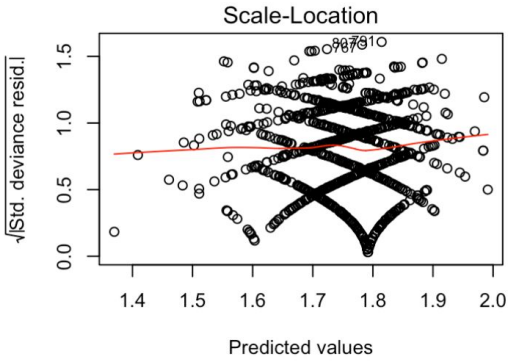
Residual Plots: Male Model

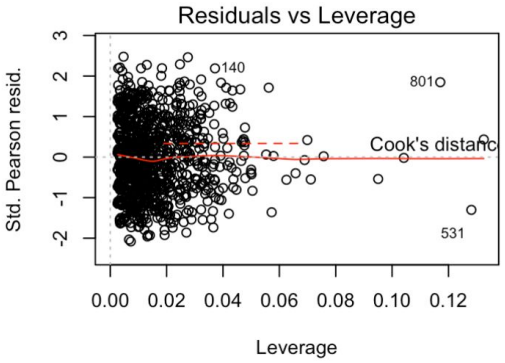


Residual Plots: Female Model









Interpretation of Results: Model 1

- Statistically significant factors: “momsedu”, “mathealth_count”, “sex“, “deaf1”
- A level increase in the mother’s education will increase the risk score by a factor of 1.02.
- Each additional mental health diagnosis for the mother will decrease the risk score by a factor of 0.3.
- If a child is male, then the risk score will be increased by a factor of 1.04.
- If a child has more than one deaf family member, then the child’s risk score will increase by a factor of 1.07.

6.1.1 Predata Table

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.6618	0.0588	28.27	0.0000
momsdeliveryage	0.0010	0.0019	0.54	0.5866
momsedu	0.0208	0.0092	2.27	0.0230
delivery2	-0.0363	0.0305	-1.19	0.2342
delivery3	0.0354	0.0239	1.48	0.1381
delivery4	0.0093	0.0238	0.39	0.6964
babymortality	-0.0123	0.0112	-1.09	0.2745
mathealth_count	-0.0289	0.0095	-3.04	0.0023
sex1	0.0411	0.0175	2.34	0.0191
tri_alcohol	0.0147	0.0268	0.55	0.5833
tri_tobacco	-0.0035	0.0241	-0.15	0.8831
mom_trauma1	-0.0145	0.0265	-0.55	0.5827
maternalpregnancyproblems____81	0.0012	0.0282	0.04	0.9654
artmethod____11	-0.0407	0.0477	-0.85	0.3939
tri_vaping	-0.0255	0.0396	-0.64	0.5208
deaf1	0.0673	0.0284	2.37	0.0177
tri_cannabis	-0.0136	0.0240	-0.57	0.5711

Interpretation of Results: Model 2

- Statistically significant factors: “mathealthcount”, “deaf1”, “maternalpregnancyproblems__8”
- An increase in mental health diagnoses for the mother decreases their child’s risk score by about 0.03.
- If the child has more than one deaf family member, their risk score increases by a factor of 1.09.
- If a male child has a mother that has anemia, then their risk score will decrease by a factor of about 0.8.
- In comparison to our first model, we observe that the variable for mother’s education is no longer significant with a p-value calculated of 0.6389.

6.1.3 Maledata Table

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.7743	0.0787	22.56	0.0000
momsdeliveryage	0.0011	0.0026	0.44	0.6574
momsedu	0.0063	0.0125	0.51	0.6123
delivery2	-0.0232	0.0409	-0.57	0.5706
delivery3	0.0192	0.0332	0.58	0.5636
delivery4	-0.0334	0.0323	-1.03	0.3010
babymortality	-0.0072	0.0148	-0.49	0.6273
mathealth_count	-0.0268	0.0130	-2.07	0.0389
tri_alcohol	-0.0268	0.0388	-0.69	0.4893
tri_tobacco	0.0405	0.0310	1.30	0.1922
mom_trauma1	-0.0560	0.0351	-1.59	0.1110
maternalpregnancyproblems____81	-0.0807	0.0414	-1.95	0.0511
artmethod____11	-0.0742	0.0729	-1.02	0.3084
tri_vaping	-0.0834	0.0578	-1.44	0.1488
deaf1	0.0859	0.0386	2.23	0.0260
tri_cannabis	-0.0161	0.0353	-0.46	0.6486

Interpretation of Results: Model 3

- Statistically significant factors: “momsedu”, “delivery3”, “delivery4”., “mathealth_count”, “tri_tobacco”, “maternalpregnancyproblems__8”
- A level increase in the mother’s education will increase a child’s risk score by a factor of 1.04.
- Delivery through a planned C-section increases the risk score by a factor of 1.06. Emergency C-section increases this score by a factor of 1.07.
- An increase in the number of mental health diagnoses for the mother decrease their child’s risk score by about 0.3.
- An increase in the number of trimesters that a mother used tobacco decreases their child’s risk score by 0.7.
- If a mother had anemia during their pregnancy, their child’s risk score will increase by 1.1.

6.1.5 Femaledata Table

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.5683	0.0852	18.41	0.0000
momsdeliveryage	0.0009	0.0028	0.34	0.7351
momsedu	0.0378	0.0136	2.79	0.0053
delivery2	-0.0548	0.0461	-1.19	0.2346
delivery3	0.0598	0.0345	1.73	0.0833
delivery4	0.0629	0.0354	1.78	0.0757
babymortality	-0.0160	0.0173	-0.93	0.3530
mathealth_count	-0.0287	0.0139	-2.07	0.0382
tri_alcohol	0.0561	0.0374	1.50	0.1335
tri_tobacco	-0.0720	0.0380	-1.90	0.0580
mom_trauma1	0.0445	0.0404	1.10	0.2717
maternalpregnancyproblems____81	0.0924	0.0388	2.38	0.0171
artmethod____11	-0.0185	0.0638	-0.29	0.7717
tri_vaping	0.0464	0.0543	0.85	0.3932
deaf1	0.0493	0.0424	1.16	0.2450
tri_cannabis	-0.0099	0.0333	-0.30	0.7673

Conclusions

- Models point to diverging risk factors for ASD behaviors between males and female
  - significant difference when looking at predictors for autism in males and females
- Significant in all models:
  - number of mental health diagnoses for the mother
- Significant in two models:
  - whether or not the mother had anemia during pregnancy
  - whether the child has more than one deaf family member
  - mother's education level
- Significant in one model:
  - two of the delivery methods of the pregnancy (those of which being planned and emergency C-section)
  - the number of trimesters a mother used tobacco

Solvable Challenges

Data Cleaning

- Solution for NAs: removed observation or set to NAs to 0
- Solution for `other` columns: categorize into columns, grouped to form new column, removed
- Solution for outliers: outliers with lots of NAs or inconsistencies were removed

Modeling

- Solution for column type: convert binary columns and non-numeric variables to factors

Unsolvable Challenges

- Low test prediction accuracy with prediction range for outcome variable of 5 to 7
- Variables we expected to be risk factors for ASD were not statistically significant
  - `asd' (number of family members with ASD)
  - ‘mat_pssd` (maternal psychosis, schizophrenia, or schizoaffective disorder)
- Small number of statistically significant predictors

Recommendations for Future Work

- Explore postnatal risk factors for ASD
- Additional research on delivery method and its effect on ASD
    - Potentially explore if any reasons for having a cesarean section (prolonged labor, abnormal positioning, cord prolapse, etc.) increase the risk score for ASD
- Condense maternal health problems on survey to include specific categories in order to avoid low column counts that skew the data
- Condense similar questions on survey to avoid collinearity

Works Cited

"Autism and Gender." *Autism Support - Leading UK Charity - National Autistic Society*, www.autism.org.uk/about/what-is/gender.aspx.

"Autism spectrum disorder." *Mayo Clinic*, 6 Jan. 2018, www.mayoclinic.org/diseases-conditions/autism-spectrum-disorder/symptoms-causes/syc-20352928.

Geelhand, Philippine. "The Role of Gender in the Perception of Autism Symptom Severity and Future Behavioral Development." *Molecular Autism*, 2019, molecularautism.biomedcentral.com.

Halladay, Alycia K, et al. "Sex and Gender Differences in Autism Spectrum Disorder: Summarizing Evidence Gaps and Identifying Emerging Areas of Priority." *Molecular Autism*, BioMed Central, 13 June 2015, www.ncbi.nlm.nih.gov/pmc/articles/PMC4465158/.

"Poisson Regression Analysis Using SPSS Statistics." *How to Perform a Poisson Regression Analysis in SPSS Statistics | Laerd Statistics*, statistics.laerd.com/spss-tutorials/poisson-regression-using-spss-statistics.php.

Vatcheva, Kristina P, et al. "Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies." *Epidemiology (Sunnyvale, Calif.)*, U.S. National Library of Medicine, Apr. 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC4888898/.