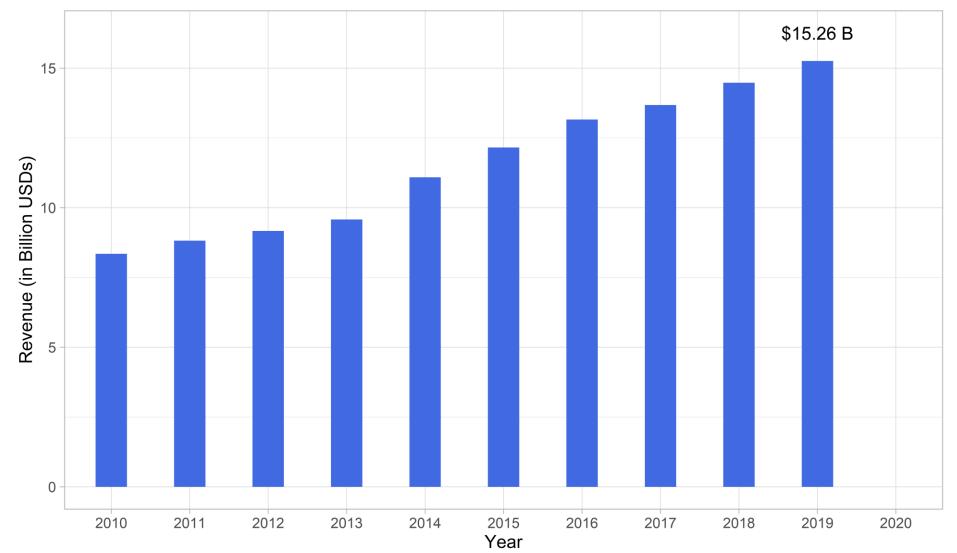
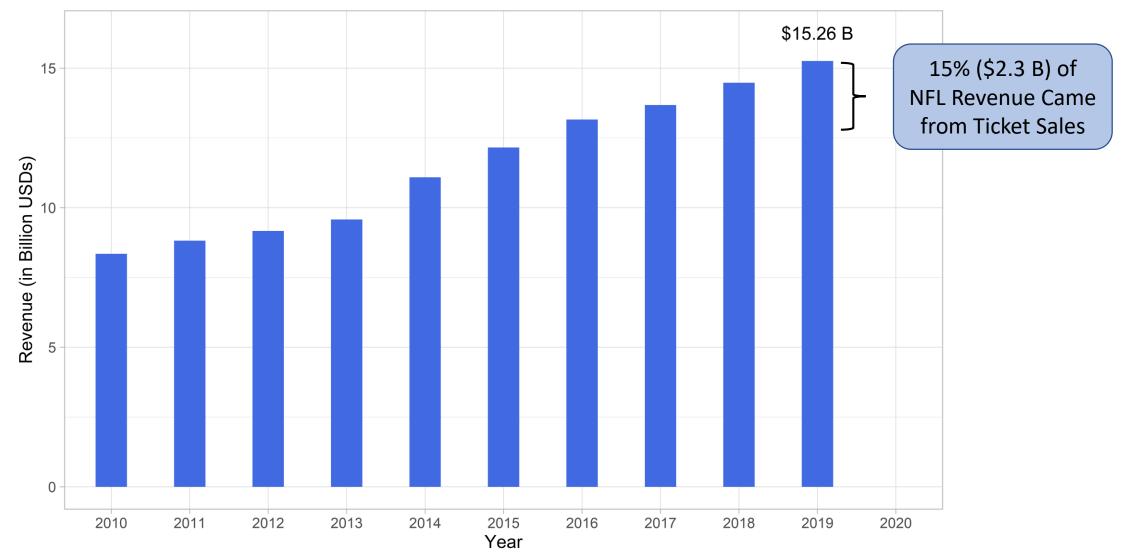
NFL Attendance Using Bayesian Hierarchical Time Series

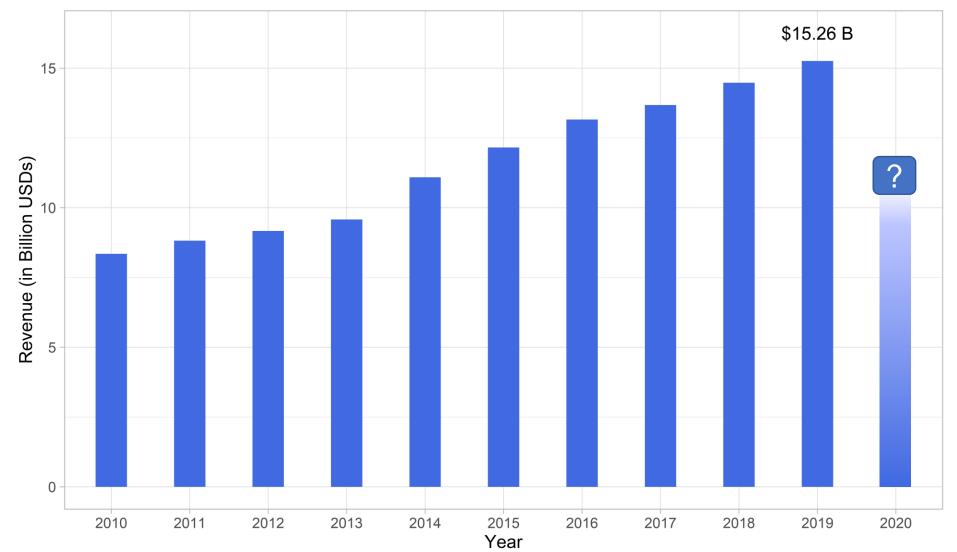
Cason Wight



Total NFL Revenue



Total NFL Revenue



Total NFL Revenue

Win Percentage 0% 20% 40% 60% 80% Patriots Steelers Packers Colts Eagles Ravens Seahawks Saints Broncos Cowboys Chargers Falcons Chiefs Vikings Team Bears Panthers Titans Giants 49ers Bengals Jets Dolphins Rams Cardinals Bills **Buccaneers** Redskins Jaguars Raiders Lions Browns 50 100 150 200 0 **Total Games Won**

Regular Game Wins (2000 to 2019)

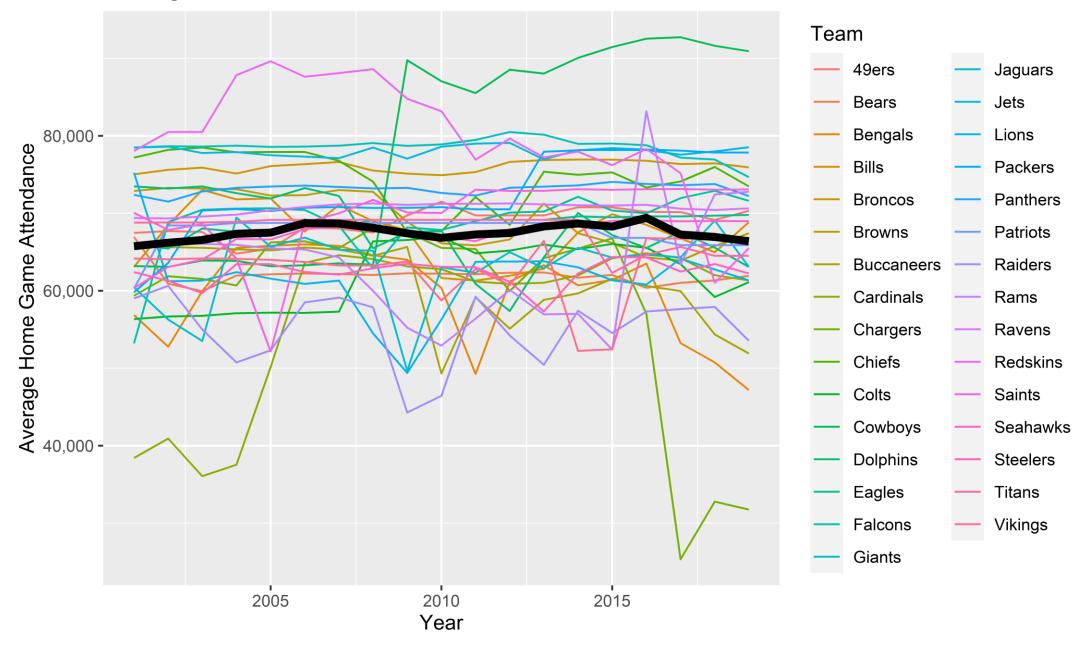
Tidy Tuesday Project

Data

31 Included Teams

16 ordinary games per year

Average Home Attendance for each Team



Research Questions

How much does one year's attendance affect the next (for each team)?

Do winning teams have higher attendance the following year?

How much ticket sales money has been lost in 2020?

Model Structure

$$y_{i,t} | \alpha_i, \beta_i, y_{i,t-1}, \theta, x_{i,t}, \sigma \sim \mathcal{N} \left(\alpha_i + \beta_i * y_{i,t-1} + \theta * x_{i,t}, \sigma^2 \right) \\ \alpha_i | \alpha \sim \mathcal{N} (\alpha, \lambda^2) \\ \alpha \sim \mathcal{N} (\mu_\alpha, \sigma_\alpha^2) \\ \beta_i | \beta \sim \mathcal{N} (\beta, \eta^2) \\ \beta \sim \mathcal{N} (\mu_\beta, \sigma_\beta^2) \\ \theta \sim \mathcal{N} (\mu_\theta, \sigma_\theta^2) \\ \sigma \sim \text{Gamma}(\alpha_\sigma, \beta_\sigma)$$

Computational Methodology

Stan (Hamiltonian MCMC)

4 chains, warmup of 1,000, thinning by 2 Takes roughly 1.2 hrs, using 4 cores for 36,000 samples

By hand (Metropolis Algorithm with MVN proposals) 1 chain (for comparison), warmup of 5,000, thinning by 25 Takes roughly 14.5 hrs, using 1 core for 20,000 samples

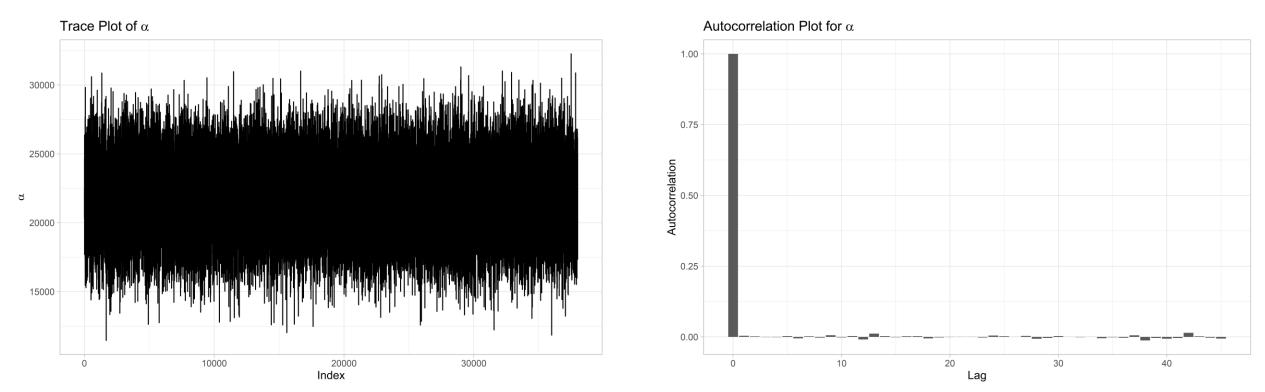
Frequentist Analysis

Simple Linear model $y_{i,t} = \alpha_i + \beta_i y_{i,t-1} + \theta x_{i,t} + \epsilon_{i,t}, \ \epsilon_{i,t} \sim \mathcal{N}(0,\sigma)$

Prior Selection

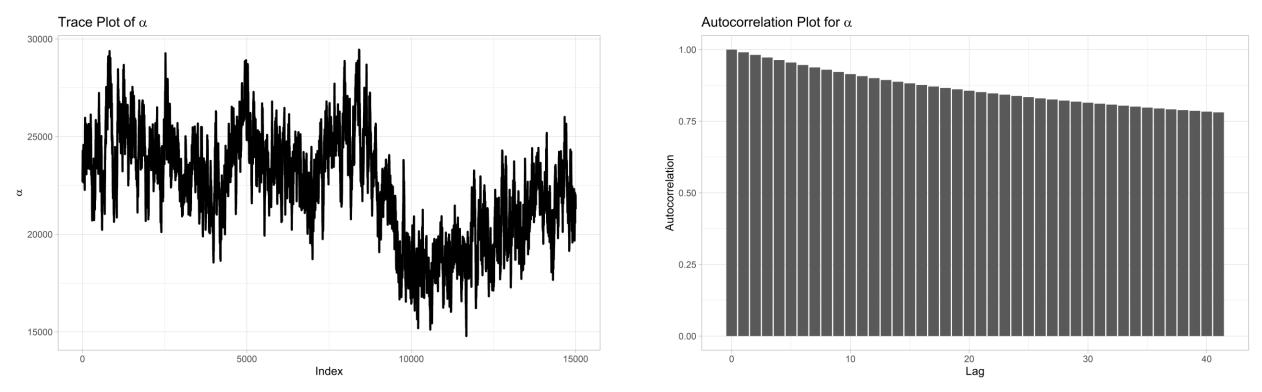
Parameter	Explanation	Prior 1	Prior 2	Prior 3
μ_{α}	Center of overall attendance intercept	30,000	66,000	66,000
σ_{lpha}	Std. deviation of overall attendance intercept	20,000	2,000,000	6,000
λ	Std. deviation of individual attendance intercepts from center	8,000	100,000	15,000
$\mu_{oldsymbol{eta}}$	Center of overall prior attendance effect	.7	0	.6
σ_{meta}	Standard deviation of overall prior attendance effect	.5	2,000,000	.2
η	Std. deviation of individual prior attendance effects from center	.1	20	.4
$\mu_{ heta}$	Center of prior wins effect	1,000	0	1,000
$\sigma_{ heta}$	Standard deviation of prior wins effect	1,000	2,000,000	500
$lpha_{\sigma}$	Shape of error in model $\frac{\text{mean}^2}{\text{variance}}$	$\left(\frac{10,000}{15,000}\right)^2$	$\left(\frac{5,000}{100,000}\right)^2$	$\left(\frac{5,000}{1,000}\right)^2$
eta_σ	Rate of error in model $\frac{\text{mean}}{\text{variance}}$	$\frac{10,000}{15,000^2}$	$\frac{5,000}{100,000^2}$	$\frac{5,000}{1,000^2}$

Diagnostics (stan results for α)



Effective sample size of 36,000 (> 30,000 for all parameters) \hat{R} of .9999 (roughly 1 for all parameters)

Diagnostics (by hand results for α)



Effective sample size of 48 (6 to 1,920 across different parameters)

Comparison of Models

Parameter	Prior 1 (stan)	Prior 1 (by hand)	Prior 2 (stan)	Prior 3 (stan)	LM
θ	269.17	271.53 (±4.10)	310.06	297.05	314.09
α	21,823.86	22,440.22 (±378.42)	28,335.97	36,863.83	—
β	.64	.63 (±.01)	.48	.51	_
σ	3,683.37	3,677.55 (±2.54)	3,626.81	3,635.34	3,631.00

Team Effect of Previous Attendance **Team Attendance Intercepts** Vikings Titans Steelers Seahawks Saints Vikings Titans Steelers Seahawks Saints Redskins Redskins Ravens Ravens Rams Raiders Raiders Patriots Patriots Panthers Model Model Panthers Packers Packers Lions Lions LM Jets Jets Team Team Jaguars Giants Falcons Eagles Dolphins Cowboys Colts Chiefs Jaguars _Giants Prior 1 Falcons Eagles Dolphins Cowboys Colts Chiefs Prior 2 Prior 3 Chargers Cardinals Chargers Cardinals Buccaneers Buccaneers Browns Broncos Bills Browns Broncos Bills Bengals Bears 49ers Bengals Bears 49ers 0.0 0.5 50000 75000 0 25000 Effect Intercept

LM

Prior 1

Prior 2

Prior 3

Bayesian Results

Prior 1 (tight priors)

Effect of Previous Year's Attendance Effect of Previous Year's Attendance Estimated Attendance 00000 00000 00000 Estimated Attendance 00000 00000 00000 Estimated Attendance 00000 00000 00000 40000 40000 40000 60000 50000 60000 70000 80000 50000 70000 80000 Prior Year Attendance **Prior Year Attendance** Density for θ Density for θ Density 0.006 Density 0.004 0.002 Density 0.006 Density 0.004 0.002 **Density** 0.006 **Density** 0.002 0.000 0.000 0.000 200 300 400 500 400 600 100 200 0 θ θ Density for σ Density for σ Density 0.003 0.002 0.001 Density 0.003 0.002 0.001 Density 0.003 0.002 0.001 0.000 0.000 0.000 3250 3500 3750 4000 4250 3250 3500 3750 4000 σ σ

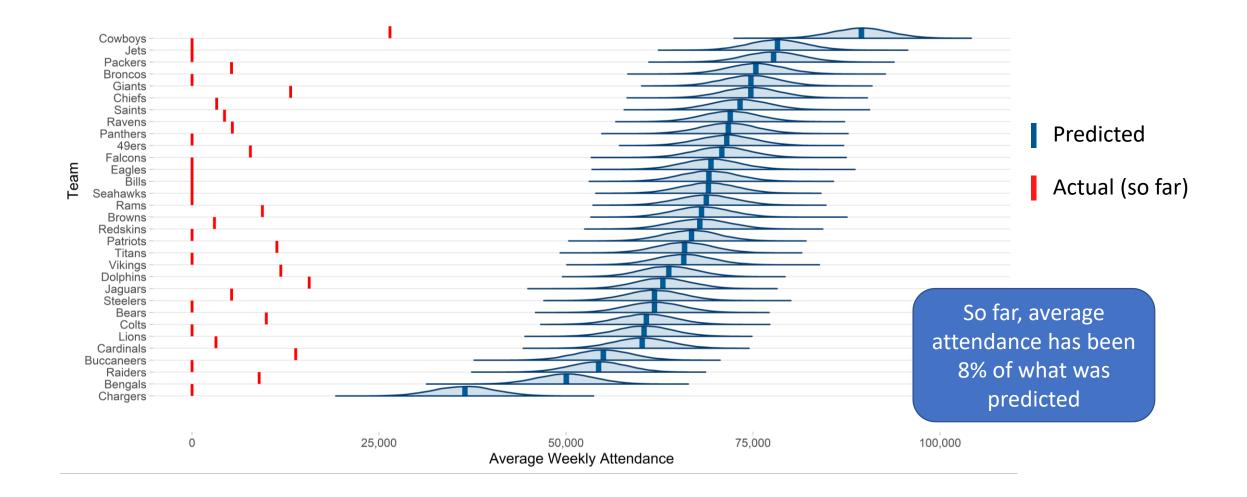
Prior 2 (noninformative priors)

Effect of Previous Year's Attendance 50000 60000 80000 70000 Prior Year Attendance Density for θ 200 100 300 400 500 θ Density for σ 3200 3400 3600 3800 4000 420

σ

Prior 3 (some tight, some not)

2020 Projections



Conclusions

Lag(1) attendance effect is largest for Cowboys, Chargers, and Cardinals

Lag(1) attendance effect is smallest for Rams, Jaguars, and Raiders

Attendance this year is down 92% from predictions

- Roughly \$2.12 Bil. lost (previous year's dollar/seat * seats lost in 2020)
- Jets, Packers, and Giants most affected so far (most seats lost)
- Chargers, Bengals, and Buccaneers least affected (least seats lost)

Each win adds roughly 300 attendees home games the next year

A simple AR(1) linear model arrives at essentially the same conclusions

Further Research

The effect of games θ and variance σ^2 could also be team-specific

Other covariates (e.g. weather or opposing team) may be interesting

Team-specific variables could be clustered