

Models

16

- **16 Nonlinear Regression Models**

- 2 Weibull Distribution Models
- 14 Shifted-Gompertz Distribution Models
 - Twelve come from running BY statement for Mkt_Genre to model for each movie genre

12

- **12 Quantile Regression Model**

- Twelve come from running BY statement for Mkt_Genre to model for each movie genre

28

- **28 Overall Models**

Models Overview

Nonlinear Regression Using the Weibull Distribution and Shifted-Gompertz Distribution

- Fit data to a distribution for forecasting or data imputation of missing values.
- Use ordinary least squares estimation with the Marquardt-Levenberg minimization method for curve fitting.
- SAS Program
 - SAS Procedure: PROC MODEL
- R Program
 - R Packages for modeling: nlstools, nlsLM(), nlshelper
 - Supplements: minpack.lm, WeibullR, growthmodels, extraDistr and diffusion, car
 - Data Manipulation and Visualization: tidyverse, lubridate, ggplot2

Models Overview

Quantile Regression

- Extends ordinary least squares regression to better fit data under data conditions of skewness, outliers and unnormal data.
 - Models a regression line not based on a mean but a quantile such as the median at the 50th quantile.
- SAS Program
 - SAS Procedure: PROC QUANTREG
- R Program
 - R Packages for modeling: quantreg for quantile regression analysis
 - Data Manipulation and Visualization: tidyverse, lubridate, ggplot2

Transformations

From PROC EXPAND to Transforming a Time Series in R

- Transforming the intervals or frequencies of a time series into another chosen interval or frequency while accounting for any missing values caused by the transformation by interpolating missing values.
- SAS Program
 - SAS Procedure: PROC EXPAND that uses cubic spline interpolation
- R Program
 - R Packages
 - `approx()` for linear interpolation
 - `splinefun()` for cubic spline interpolation

Outputs of the SAS scripts

Estimates, Modeling, SQL Transformations

- Model information including parameters, parameter estimates, exogenous and endogenous variables, SSE and RMSE

| Model Summary | |
|----------------------|----|
| Model Variables | 2 |
| Endogenous | 1 |
| Exogenous | 1 |
| Parameters | 4 |
| OutVars Variables | 1 |
| Equations | 1 |
| Number of Statements | 14 |

| | |
|-------------------|--------------------------------|
| Model Variables | t index_act |
| Parameters(Value) | M(10000) b(0.4) eta(10) gstart |
| Equations | index_act = |

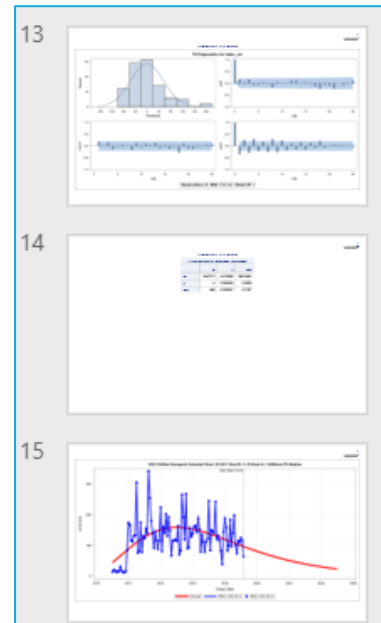
| | |
|-----------------------------|--------------|
| The Equation to Estimate is | |
| index_act = | F(M, b, eta) |

| Nonlinear OLS Summary of Residual Errors | | | | | | | |
|--|----------|----------|---------|--------|----------|----------|----------|
| Equation | DF Model | DF Error | SSE | MSE | Root MSE | R-Square | Adj R-Sq |
| index_act | 3 | 97 | 42.9236 | 0.4425 | 0.6652 | 0.8100 | 0.8061 |

| Nonlinear OLS Parameter Estimates | | | | |
|-----------------------------------|----------|----------------|---------|----------------|
| Parameter | Estimate | Approx Std Err | t Value | Approx Pr > t |
| M | 783.3232 | 169.2 | 4.63 | <.0001 |
| b | 0.014523 | 0.00223 | 6.50 | <.0001 |
| eta | 3.977083 | 0.3837 | 10.34 | <.0001 |

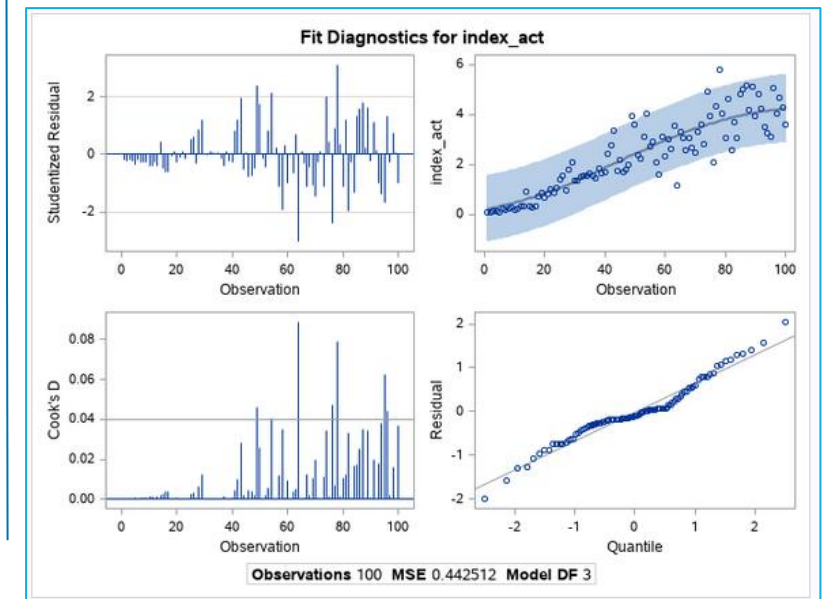
PPT

- SAS output written to PowerPoint slides.



Diagnostic

- Diagnostic plots of residuals, outliers, normality and autocorrelation.



R and SAS Scripts.

Estimates, Modeling, SQL Transformations

- Model information including parameters, parameter estimates, variables, SSE and RMSE

```
-----
Formula: MED_EST_BI_8 ~ model_fn(m, b, eta, t)

Parameters:
      Estimate Std. Error t value Pr(>|t|)
m  7.829e+02  1.689e+02   4.635 1.11e-05 ***
b  1.453e-02  2.232e-03   6.509 3.34e-09 ***
eta 3.976e+00  3.635e-01  10.938 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

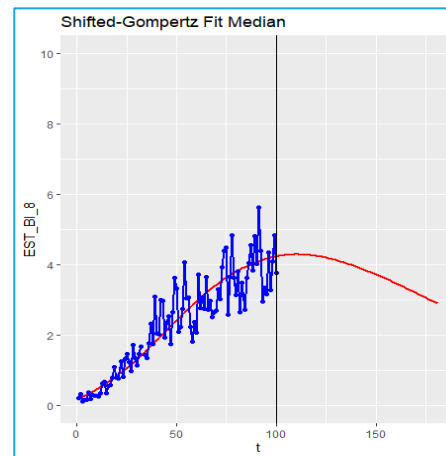
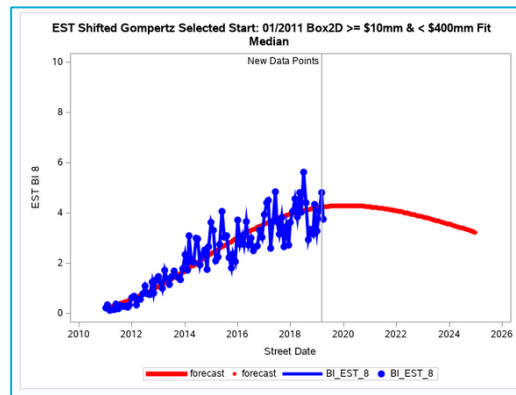
Residual standard error: 0.6651 on 97 degrees of freedom

Number of iterations to convergence: 11
Achieved convergence tolerance: 1.49e-08

-----
Residual sum of squares: 42.9
```

Growth Curve

- SAS Growth Curve and R ggplot2



Diagnostic

- Diagnostic plots of residuals, outliers, normality and autocorrelation.

