MBA 652A STATISTICAL MODELLING FOR BUSINESS ANALYTICS

PROJECT REPORT ON

DAIRY FARM MILK PRODUCTION

SUBMITTED TO:

Dr. DEVLINA CHATTERJEE

SUBMITTED BY:

Ms. ANUSHRI SINGH (19125010)

Mr. ASHWIN BHIDE (19125015)

Mr. FARAZ AHMAD KHAN (19114006)

Mr. YASH KALPESH PANCHAL (19114020)

OBJECTIVE:

The objective of the project is to determine the amount of milk each farm produce based on parameters like number of cows, Amount of feed, land area of farm, etc. by applying different regression methods available for panel data.

METHODOLOGY:

We have made Pooled model, Entity fixed(entity demeaned) model, Entity fixed(n-1 binary variable) model, Time fixed(time demeaned), Time fixed(n-1 binary variable), Entity and time both fixed model, Random effects model and performed various tests to conclude the model that best explain the amount of milk production.

DATA:

Source:

http://people.stern.nyu.edu/wgreene/Econometrics/dairy.csv

Description:

This is a Panel Data set, consisting data of Spanish Dairy Farm Production. A panel data set, also called longitudinal data set, is one that studies the same parameters at different points in time. Data has 247 different Farms and the data is taken over 6 years, therefore n = 247 and T = 6. The balance in the dataset is established as there are no missing values. Variables in the Data Set are number of cows in the farm, how much they are feed in total, land area of the farm, number of labors involved in the process, these are independent variables. While the dependent variable is the amount of Milk produced by each farm in each year.

Variables Involved in Analysis:

Sr. No.	Variable	Description
1	FARM	farm number
2	YEAR	Year
3	COWS	number of cows (mean deviation)
4	LAND	land area (in some units)
5	FEED	Amount of food fed to cows
6	LABOR	number of labors at farm (mean deviation)
7	MILK	farm Output (amount of milk produced)

Independent and Dependent Variables:

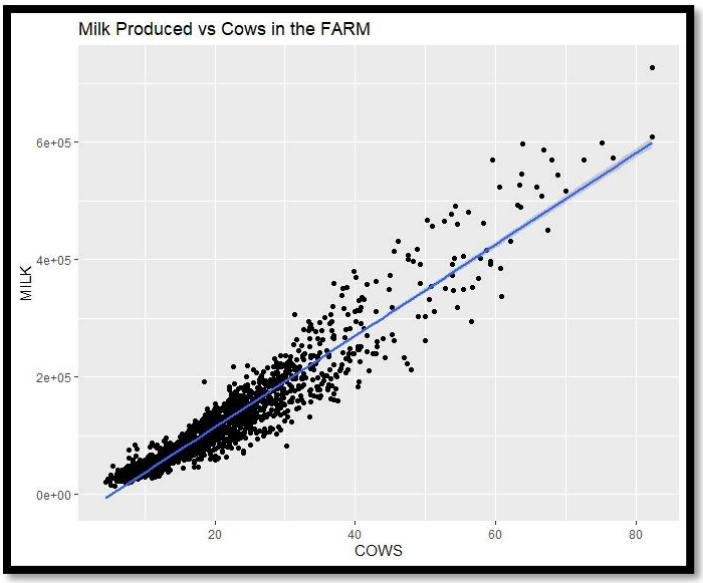
Independent Variable: COWS, LAND, FEED, LABOR

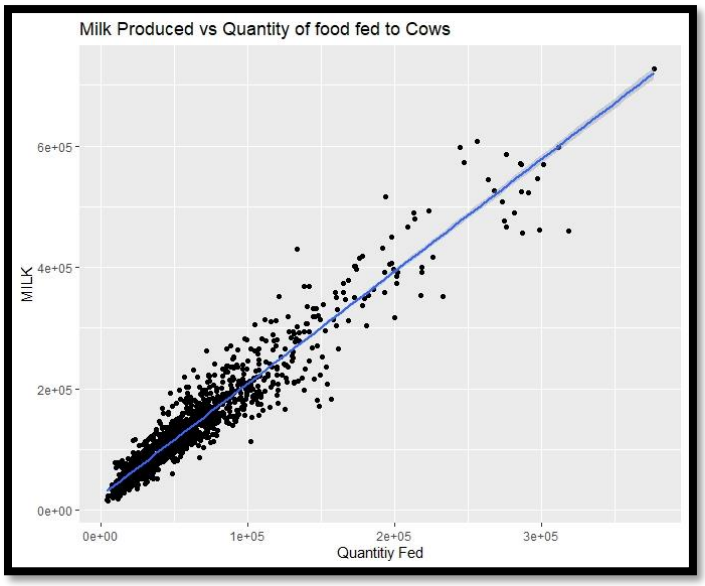
Analysis is done for which independent value is changed and the effect on dependent variable is observed or stud

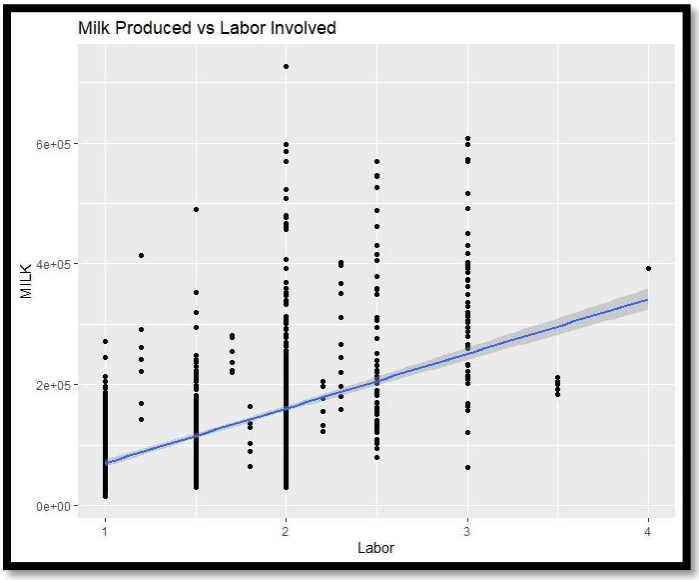
Dependent Variable: MILK

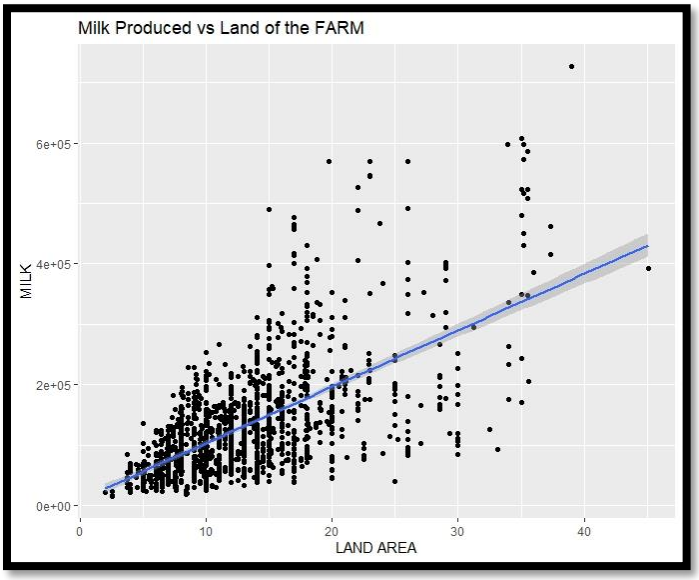
Amount of milk produced at a particular farm in each year. This is the variable which is tested and changing independent variables will affect the outcome this.

Scatter Plots:





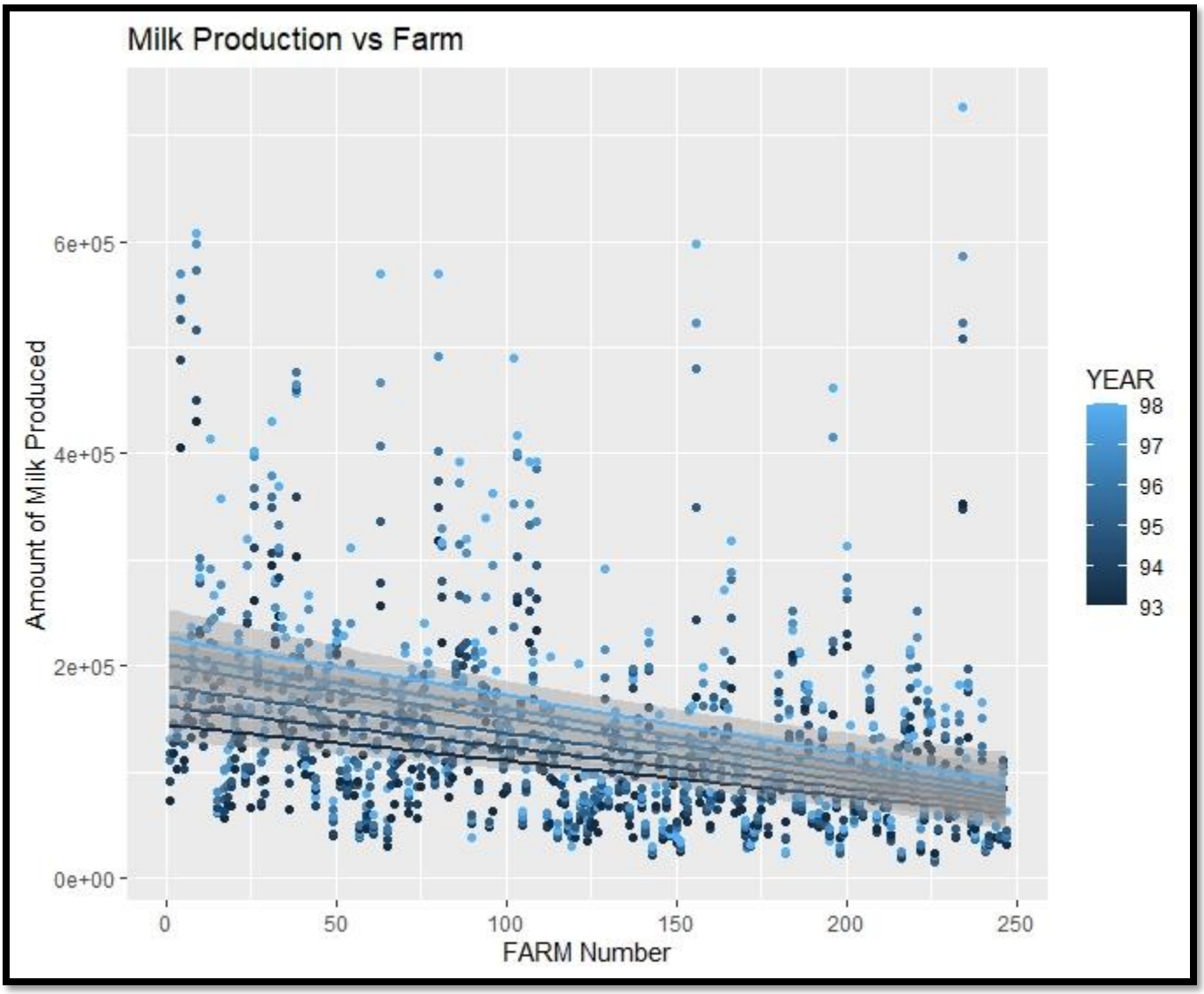


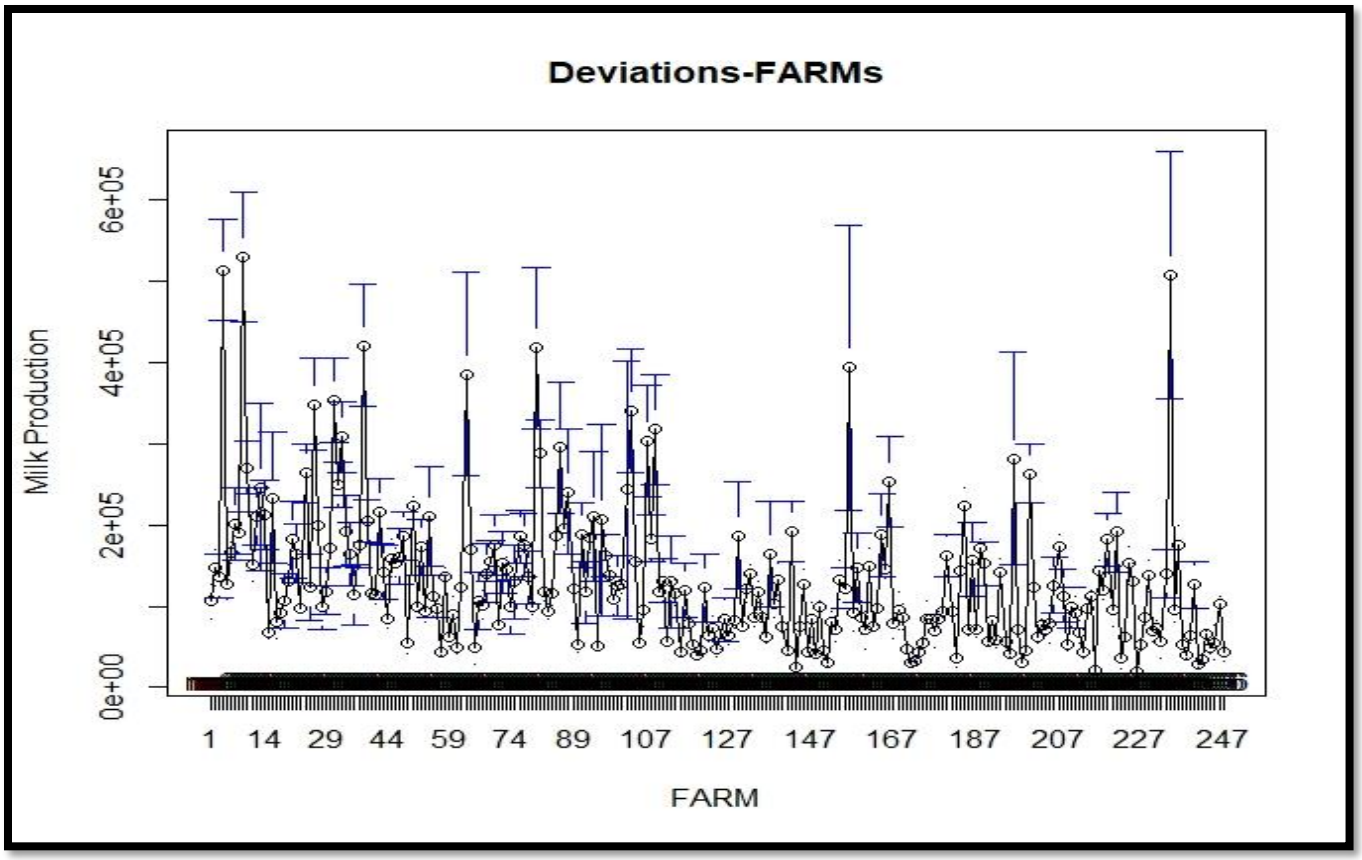


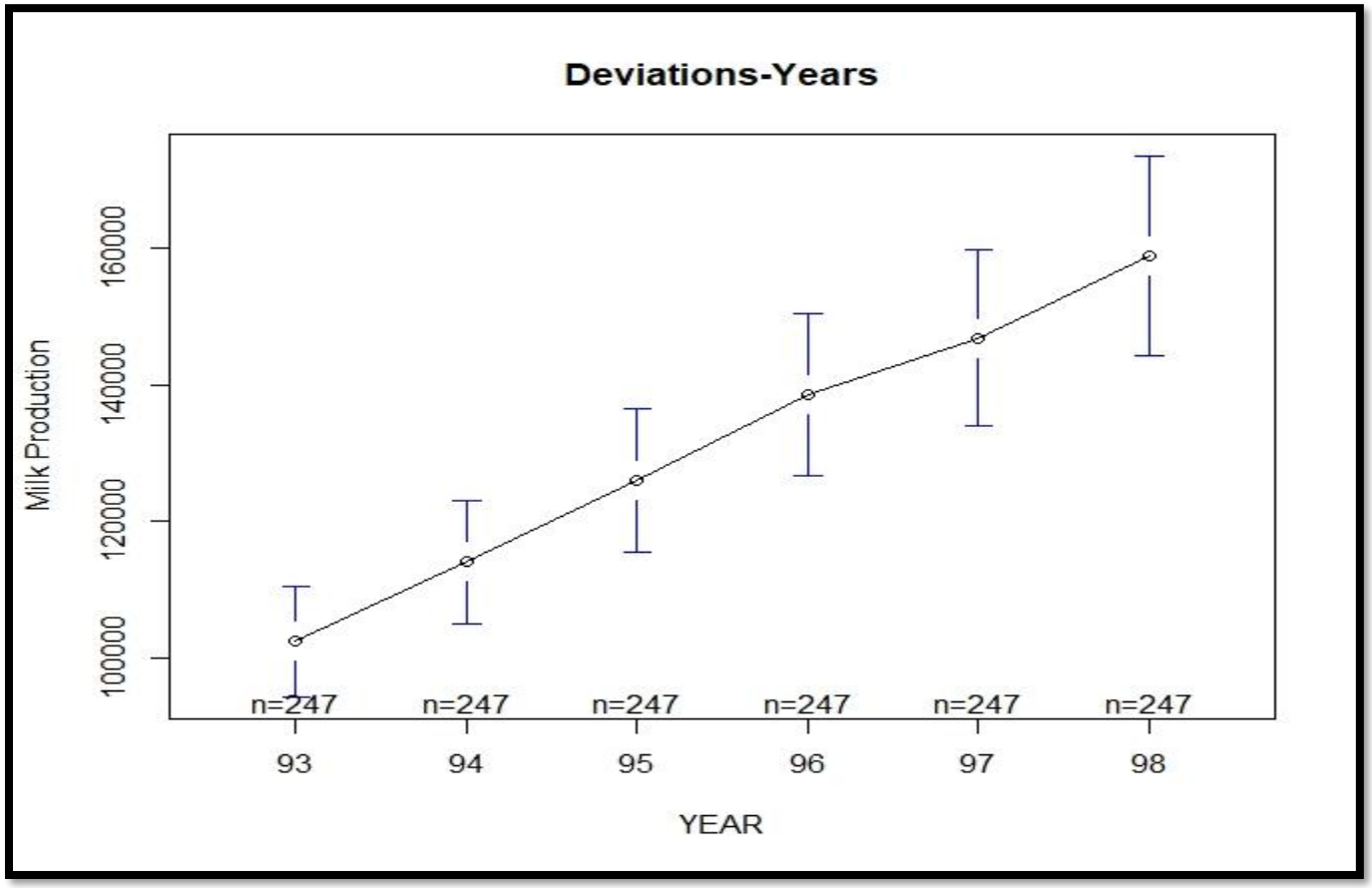
Statistical Data of Variables:

Variable	Mean	Standard Deviation	Min	Max
COWS	22.12	11.27	4.5	82.3
LAND	12.99	6.17	2	45.1
FEED	57941	47981.4	3924	376732
LABOR	1.672	0.55	1	4
MILK	131107	92583.98	14410	727281

Data Visualization:




Deviations-FARMs
Milk Production
0e+00
2e+05
4e+05
6e+05
1
14
29
44
59
74
89
107
127
147
167
187
207
227
247
FARM


Deviations-Years
Milk Production
100000
120000
140000
160000
n=247
n=247
n=247
n=247
n=247
n=247
93
94
95
96
97
98
YEAR

Pooling Model:

The pooled regression model ignores any differences over entities or time and treats each data as a separate entity. We use the "plm" function to carry out pooled regression on the data in R.

Best fit Pooling model:

```
plm(MILK ~ COWS + FEED, data=Pooled_Data, model = 'pooling')
```

Summary:

```
Balanced Panel: n = 247, T = 6, N = 1482

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-106439.79   -9455.23     397.36    9541.44   86354.93

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
(Intercept) -1.2093e+04  1.3075e+03 -9.2493 < 2.2e-16 ***
COWS         3.6665e+03  1.0455e+02 35.0694 < 2.2e-16 ***
FEED         1.0720e+00  2.4566e-02 43.6366 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    1.2695e+13
Residual Sum of Squares: 5.7666e+11
R-Squared:      0.95458
Adj. R-Squared: 0.95451
F-statistic: 15540.2 on 2 and 1479 DF, p-value: < 2.22e-16
```

Observation:

Adjusted R^2 value is quite high, meaning that only 2 variables COWS and FEED are enough to explain the dependent variable MILK, but we will look for other models, which are Time Fixed and Entity Fixed.

Entity Fixed Model (Entity Demeaned):

In entity fixed models, each FARM is considered as a separate entity. In entity demeaned method, all the values across the years, for each FARM are averaged and the average is then subtracted from the each observation. This way, we obtain the entity demeaned observations which are then regressed.

Best fit Model:

```
plm(MILK ~ COWS + FEED, data=Pooled_Data, model = 'within', effect='individual')
```

Summary:

```
Balanced Panel: n = 247, T = 6, N = 1482

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-83197.816  -4620.272     47.984   4842.493  63669.475

Coefficients:
      Estimate Std. Error t-value  Pr(>|t|)
COWS 4.1634e+03 1.7011e+02  24.476 < 2.2e-16 ***
FEED 9.5899e-01 2.6762e-02  35.834 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Total Sum of Squares:    1.512e+12
Residual Sum of Squares: 1.8806e+11
R-Squared:      0.87562
Adj. R-Squared: 0.8506
F-statistic: 4340.05 on 2 and 1233 DF, p-value: < 2.22e-16
```

Observation:

This model too has quite high Adjusted R^2 value, and the same 2 variables are used to explain dependent variable, as other independent variables are not significant. We would look for other models.

Entity Fixed Model ((n-1) Binary Variable):

In this model each entity except one is assigned to a new binary variable (in total n-1 variables) which has value 1 for same entity and 0 for rest entities.

Best Fit Model:

```
lm(MILK ~ COWS + FEED + FARM, data=Pooled_Data)
```

Summary:

```
Residuals:
   Min     1Q Median     3Q    Max
-83198  -4620     48   4842  63669

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.244e+04  5.471e+03  -2.273 0.023170 *
COWS         4.163e+03  1.701e+02  24.476  < 2e-16 ***
FEED         9.590e-01  2.676e-02  35.834  < 2e-16 ***
FARM2        6.560e+03  7.152e+03   0.917 0.359222
FARM3       -8.920e+03  7.178e+03  -1.243 0.214205
FARM4        1.669e+04  8.641e+03   1.931 0.053679 .
FARM5        2.749e+03  7.134e+03   0.385 0.700029
FARM6        6.622e+03  7.157e+03   0.925 0.355004
FARM7       -5.548e+03  7.331e+03  -0.757 0.449286
FARM8        8.059e+02  7.219e+03   0.112 0.911131
FARM9        3.753e+04  9.604e+03   3.908 9.81e-05 ***
FARM10       2.859e+04  7.336e+03   3.897 0.000102 ***
FARM11       1.741e+04  7.153e+03   2.433 0.015096 *
FARM12       1.611e+04  7.316e+03   2.202 0.027842 *
FARM13       2.273e+04  7.325e+03   3.104 0.001954 **
FARM14      -4.082e+03  7.288e+03  -0.560 0.575477
FARM15      -2.442e+04  7.205e+03  -3.389 0.000725 ***
FARM16       6.353e+03  7.398e+03   0.859 0.390650
FARM17      -1.576e+04  7.132e+03  -2.209 0.027330 *
FARM18      -1.642e+03  7.132e+03  -0.230 0.817972
FARM19      -1.484e+04  7.157e+03  -2.074 0.038278 *
FARM20      -1.918e+04  7.153e+03  -2.681 0.007433 **
FARM21       1.178e+04  7.181e+03   1.640 0.101194
FARM22       3.654e+03  7.173e+03   0.509 0.610519
FARM23      -1.913e+04  7.216e+03  -2.651 0.008123 **
FARM24       5.960e+03  7.660e+03   0.778 0.436671
FARM25      -6.051e+03  7.166e+03  -0.844 0.398618
FARM26       8.664e+03  8.498e+03   1.019 0.308201
FARM27      -2.801e+04  7.566e+03  -3.701 0.000224 ***
FARM28      -1.209e+04  7.131e+03  -1.695 0.090237 .
FARM29      -8.208e+03  7.134e+03  -1.151 0.250107
FARM30       1.213e+04  7.155e+03   1.695 0.090305 .
FARM31       4.945e+04  7.463e+03   6.625 5.17e-11 ***
```

FARM32	2.210e+04	7.286e+03	3.034	0.002465 **
FARM33	3.708e+04	7.599e+03	4.879	1.21e-06 ***
FARM34	-2.725e+04	7.393e+03	-3.686	0.000237 ***
FARM35	-1.411e+04	7.441e+03	-1.895	0.058261 .
FARM36	1.008e+04	7.132e+03	1.413	0.157985
FARM37	-2.405e+04	7.213e+03	-3.334	0.000881 ***
FARM38	-2.342e+04	8.019e+03	-2.920	0.003562 **
FARM39	-2.342e+04	7.252e+03	-3.229	0.001275 **
FARM40	-9.060e+03	7.134e+03	-1.270	0.204347
FARM41	-2.453e+04	7.144e+03	-3.434	0.000616 ***
FARM42	-1.969e+04	7.272e+03	-2.708	0.006854 **
FARM43	-1.342e+04	7.154e+03	-1.875	0.060990 .
FARM44	-1.754e+04	7.140e+03	-2.456	0.014171 *
FARM45	-5.113e+03	7.164e+03	-0.714	0.475492
FARM46	-2.059e+04	7.224e+03	-2.850	0.004448 **
FARM47	-6.560e+03	7.170e+03	-0.915	0.360407
FARM48	-1.582e+03	7.261e+03	-0.218	0.827550
FARM49	-5.881e+03	7.149e+03	-0.823	0.410862
FARM50	-5.348e+04	7.585e+03	-7.050	2.97e-12 ***
FARM51	-1.666e+04	7.131e+03	-2.336	0.019656 *
FARM52	-1.199e+04	7.233e+03	-1.657	0.097738 .
FARM53	-5.231e+03	7.134e+03	-0.733	0.463583
FARM54	-2.872e+04	7.441e+03	-3.860	0.000119 ***
FARM55	-1.058e+04	7.158e+03	-1.477	0.139840
FARM56	-2.134e+04	7.150e+03	-2.985	0.002890 **
FARM57	6.784e+03	7.232e+03	0.938	0.348385
FARM58	-6.123e+02	7.142e+03	-0.086	0.931692
FARM59	-1.053e+04	7.148e+03	-1.474	0.140863
FARM60	-1.143e+04	7.135e+03	-1.601	0.109563
FARM61	8.585e+03	7.214e+03	1.190	0.234267
FARM62	-7.106e+03	7.141e+03	-0.995	0.319919
FARM63	3.662e+04	7.686e+03	4.765	2.11e-06 ***
FARM64	-3.914e+04	7.240e+03	-5.407	7.70e-08 ***
FARM65	-5.077e+03	7.180e+03	-0.707	0.479593
FARM66	-2.556e+04	7.147e+03	-3.576	0.000362 ***
FARM67	-7.600e+03	7.131e+03	-1.066	0.286724
FARM68	-1.206e+04	7.149e+03	-1.687	0.091897 .
FARM69	-7.018e+04	7.253e+03	-9.675	< 2e-16 ***
FARM70	-2.148e+04	7.209e+03	-2.980	0.002941 **
FARM71	5.024e+03	7.166e+03	0.701	0.483384
FARM72	-3.049e+04	7.251e+03	-4.205	2.80e-05 ***
FARM73	2.400e+03	7.143e+03	0.336	0.736926
FARM74	-8.832e+03	7.130e+03	-1.239	0.215712
FARM75	7.247e+03	7.172e+03	1.010	0.312470
FARM76	6.416e+03	7.216e+03	0.889	0.374106
FARM77	-1.302e+04	7.255e+03	-1.794	0.073062 .
FARM78	-5.131e+02	7.163e+03	-0.072	0.942908
FARM79	-1.967e+03	7.166e+03	-0.274	0.783801
FARM80	-7.190e+03	8.466e+03	-0.849	0.395884
FARM81	1.480e+04	7.469e+03	1.981	0.047792 *
FARM82	-4.084e+03	7.205e+03	-0.567	0.570947
FARM83	-1.702e+04	7.212e+03	-2.360	0.018408 *
FARM84	-8.059e+02	7.167e+03	-0.112	0.910492
FARM85	-2.311e+04	7.454e+03	-3.101	0.001976 **
FARM86	1.057e+04	7.566e+03	1.398	0.162496
FARM87	3.540e+04	7.156e+03	4.947	8.59e-07 ***
FARM88	2.948e+04	7.249e+03	4.067	5.07e-05 ***
FARM89	-3.316e+03	7.139e+03	-0.465	0.642326
FARM90	4.417e+03	7.189e+03	0.614	0.539092
FARM91	-6.330e+03	7.211e+03	-0.878	0.380215
FARM92	-6.507e+03	7.153e+03	-0.910	0.363116
FARM93	-2.867e+03	7.274e+03	-0.394	0.693580
FARM94	1.697e+04	7.223e+03	2.349	0.018991 *
FARM95	-6.193e+03	7.183e+03	-0.862	0.388715
FARM96	7.764e+03	7.226e+03	1.074	0.282838

FARM97	1.100e+03	7.200e+03	0.153	0.878566	
FARM98	-4.354e+02	7.207e+03	-0.060	0.951840	
FARM99	-9.952e+03	7.138e+03	-1.394	0.163478	
FARM100	6.572e+03	7.151e+03	0.919	0.358250	
FARM101	1.882e+04	7.134e+03	2.638	0.008449	**
FARM102	-1.191e+04	7.307e+03	-1.630	0.103334	
FARM103	-3.538e+03	7.798e+03	-0.454	0.650135	
FARM104	-1.298e+04	7.193e+03	-1.805	0.071380	.
FARM105	-1.815e+04	7.149e+03	-2.538	0.011257	*
FARM106	-4.774e+03	7.140e+03	-0.669	0.503837	
FARM107	-4.953e+03	7.856e+03	-0.630	0.528549	
FARM108	1.440e+04	7.249e+03	1.987	0.047195	*
FARM109	-4.972e+04	8.485e+03	-5.860	5.94e-09	***
FARM110	6.434e+03	7.133e+03	0.902	0.367273	
FARM111	-9.479e+03	7.149e+03	-1.326	0.185144	
FARM112	-7.416e+03	7.153e+03	-1.037	0.300050	
FARM113	7.482e+03	7.134e+03	1.049	0.294498	
FARM114	6.324e+03	7.153e+03	0.884	0.376822	
FARM115	6.923e+03	7.205e+03	0.961	0.336824	
FARM116	-4.404e+03	7.172e+03	-0.614	0.539284	
FARM117	-5.624e+03	7.138e+03	-0.788	0.430915	
FARM118	-7.003e+03	7.163e+03	-0.978	0.328450	
FARM119	1.423e+03	7.189e+03	0.198	0.843182	
FARM120	-5.100e+03	7.159e+03	-0.712	0.476365	
FARM121	1.327e+04	7.243e+03	1.832	0.067213	.
FARM122	3.452e+03	7.148e+03	0.483	0.629192	
FARM123	-2.034e+03	7.138e+03	-0.285	0.775735	
FARM124	-1.995e+03	7.158e+03	-0.279	0.780489	
FARM125	5.443e+03	7.147e+03	0.762	0.446447	
FARM126	-1.148e+04	7.133e+03	-1.609	0.107827	
FARM127	-1.736e+03	7.143e+03	-0.243	0.808018	
FARM128	-1.343e+04	7.143e+03	-1.881	0.060258	.
FARM129	1.445e+04	7.185e+03	2.011	0.044594	*
FARM130	-2.004e+03	7.143e+03	-0.281	0.779057	
FARM131	5.075e+03	7.131e+03	0.712	0.476777	
FARM132	6.969e+03	7.145e+03	0.975	0.329584	
FARM133	-1.118e+04	7.192e+03	-1.554	0.120442	
FARM134	2.380e+03	7.138e+03	0.333	0.738884	
FARM135	6.573e+03	7.149e+03	0.919	0.358099	
FARM136	1.194e+04	7.200e+03	1.658	0.097580	.
FARM137	-1.052e+04	7.244e+03	-1.453	0.146538	
FARM138	-2.614e+03	7.132e+03	-0.366	0.714065	
FARM139	1.008e+04	7.142e+03	1.412	0.158332	
FARM140	-7.350e+03	7.136e+03	-1.030	0.303207	
FARM141	-7.300e+03	7.156e+03	-1.020	0.307882	
FARM142	-8.165e+03	7.360e+03	-1.109	0.267488	
FARM143	2.728e+03	7.260e+03	0.376	0.707123	
FARM144	5.100e+03	7.164e+03	0.712	0.476652	
FARM145	-1.930e+04	7.259e+03	-2.659	0.007943	**
FARM146	8.464e+03	7.233e+03	1.170	0.242137	
FARM147	-8.932e+02	7.143e+03	-0.125	0.900501	
FARM148	-3.966e+03	7.164e+03	-0.554	0.579959	
FARM149	1.543e+04	7.135e+03	2.163	0.030704	*
FARM150	6.233e+03	7.194e+03	0.866	0.386489	
FARM151	2.077e+03	7.234e+03	0.287	0.774058	
FARM152	-7.080e+03	7.133e+03	-0.993	0.321138	
FARM153	-1.805e+04	7.154e+03	-2.523	0.011768	*
FARM154	-3.993e+03	7.206e+03	-0.554	0.579615	
FARM155	7.818e+03	7.132e+03	1.096	0.273221	
FARM156	1.058e+04	7.946e+03	1.332	0.183076	
FARM157	-4.842e+04	7.309e+03	-6.624	5.20e-11	***
FARM158	9.317e+03	7.156e+03	1.302	0.193174	
FARM159	-8.931e+03	7.134e+03	-1.252	0.210822	
FARM160	-3.306e+04	7.136e+03	-4.633	3.99e-06	***
FARM161	-1.361e+03	7.176e+03	-0.190	0.849653	

```
FARM162      3.385e+03  7.151e+03   0.473 0.636069
FARM163     -1.266e+04  7.188e+03  -1.761 0.078428 .
FARM164     -5.330e+04  7.629e+03  -6.986 4.62e-12 ***
FARM165     -9.171e+03  7.214e+03  -1.271 0.203865
FARM166      3.813e+04  7.298e+03   5.225 2.05e-07 ***
FARM167     -3.551e+04  7.204e+03  -4.929 9.39e-07 ***
FARM168      9.474e+02  7.132e+03   0.133 0.894351
FARM169     -8.468e+03  7.133e+03  -1.187 0.235367
FARM170     -3.167e+03  7.177e+03  -0.441 0.659160
FARM171     -1.262e+04  7.170e+03  -1.760 0.078585 .
FARM172     -8.905e+03  7.174e+03  -1.241 0.214756
FARM173     -9.314e+03  7.168e+03  -1.299 0.194055
FARM174      7.014e+03  7.182e+03   0.977 0.328981
FARM175      2.999e+03  7.137e+03   0.420 0.674428
FARM176     -1.665e+04  7.138e+03  -2.333 0.019806 *
FARM177      2.098e+04  7.223e+03   2.904 0.003750 **
FARM178      5.295e+03  7.147e+03   0.741 0.458890
FARM179      2.885e+03  7.150e+03   0.404 0.686626
FARM180      2.088e+03  7.171e+03   0.291 0.770926
FARM181     -1.956e+04  7.252e+03  -2.698 0.007081 **
FARM182      7.830e+03  7.234e+03   1.082 0.279318
FARM183     -8.642e+03  7.266e+03  -1.189 0.234507
FARM184     -5.349e+04  7.608e+03  -7.031 3.39e-12 ***
FARM185     -8.763e+02  7.149e+03  -0.123 0.902458
FARM186     -8.872e+03  7.387e+03  -1.201 0.229977
FARM187     -4.805e+03  7.137e+03  -0.673 0.500981
FARM188     -9.777e+03  7.643e+03  -1.279 0.201050
FARM189      4.585e+03  7.309e+03   0.627 0.530549
FARM190      6.673e+03  7.161e+03   0.932 0.351574
FARM191      1.232e+04  7.146e+03   1.724 0.084881 .
FARM192     -1.831e+02  7.153e+03  -0.026 0.979583
FARM193      1.209e+04  7.226e+03   1.673 0.094611 .
FARM194     -6.532e+03  7.158e+03  -0.912 0.361697
FARM195     -1.272e+04  7.152e+03  -1.779 0.075454 .
FARM196     -8.037e+03  7.413e+03  -1.084 0.278489
FARM197      2.307e+03  7.141e+03   0.323 0.746681
FARM198     -3.637e+03  7.179e+03  -0.507 0.612529
 [ reached getOption("max.print") -- omitted 49 rows ]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 12350 on 1233 degrees of freedom
Multiple R-squared:  0.9852,       Adjusted R-squared:  0.9822
F-statistic: 330.6 on 248 and 1233 DF,  p-value: < 2.2e-16
```

Observation:

This model has the highest Adjusted R^2 value off all models analyzed till now, but still we would continue our search for even better model.

Time Fixed Model (Time Demeaned):

In time fixed models, each YEAR is considered as a separate time. In time demeaned method, all the values across the farms, for each YEAR are averaged and the average is then subtracted from the each observation. This way, we obtain the time demeaned observations which are then regressed.

Best fit Model:

```
plm(MILK ~ COWS + LABOR + FEED, data=Pooled_Data, model = 'within', effect =
'time')
```

Summary:

```
Balanced Panel: n = 247, T = 6, N = 1482

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-105659.1   -9515.5     595.8    9679.2   84782.7

Coefficients:
        Estimate Std. Error t-value Pr(>|t|)
COWS  3.6437e+03 1.0952e+02 33.2705  < 2e-16 ***
LABOR 2.3259e+03 1.1238e+03  2.0696  0.03866 *
FEED  1.0566e+00 2.5178e-02 41.9637  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1.2151e+13
Residual Sum of Squares: 5.7185e+11
R-Squared:      0.95294
Adj. R-Squared: 0.95268
F-statistic: 9941.77 on 3 and 1473 DF, p-value: < 2.22e-16
```

Observation:

For this model, 3 independent variables are used to explain the dependent variable which is a change as compared to previous models. Also, Adjusted R^2 value is quite high, meaning that independent variables are explaining dependent variables quite efficiently.

Time Fixed Model ((T-1) Binary Variable):

In this model each time except one is assigned to a new binary variable (in total T-1 variables) which has value 1 for same time and 0 for rest of the times.

Best Fit Model:

```
lm(MILK ~ COWS + LABOR + FEED + YEAR, data=Pooled_Data)
```

Summary:

```
Residuals:
    Min      1Q  Median      3Q     Max
-105659   -9515     596    9679   84783

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.759e+04  2.166e+03  -8.118  9.9e-16 ***
COWS         3.644e+03  1.095e+02  33.270  < 2e-16 ***
LABOR        2.326e+03  1.124e+03   2.070  0.03866 *
FEED         1.057e+00  2.518e-02  41.964  < 2e-16 ***
YEAR94       2.097e+03  1.775e+03   1.182  0.23753
YEAR95       3.514e+03  1.785e+03   1.968  0.04924 *
YEAR96       4.323e+03  1.795e+03   2.409  0.01611 *
YEAR97       3.300e+03  1.808e+03   1.825  0.06818 .
YEAR98       4.765e+03  1.825e+03   2.611  0.00911 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 19700 on 1473 degrees of freedom
Multiple R-squared:  0.955, Adjusted R-squared:  0.9547
F-statistic:  3903 on 8 and 1473 DF,  p-value: < 2.2e-16
```

Observation:

Here 5 binary variables are added along with 3 significant at 95% level independent variables to explain dependent variable, this process yield Adjusted R^2 value to 0.9547 which is quite good.

Entity and Time both Fixed Model:

In this model, variations across both the time and entity are considered. Through this both the effects that vary across the farm but remain same over year and the effects that vary across the year and remain same across the farm are handled.

Best Fit Model:

```
plm(MILK ~ COWS + FEED, data=Pooled_Data, model='within',
index=c("FARM","YEAR"), effect ='twoways')
```

Summary:

```
Balanced Panel: n = 247, T = 6, N = 1482

Residuals:
    Min.  1st Qu.   Median  3rd Qu.     Max.
-80469.2  -4601.6    -51.3   4851.0  63227.1

Coefficients:
       Estimate Std. Error t-value  Pr(>|t|)
COWS 3978.39696  171.03815  23.260 < 2.2e-16 ***
FEED    0.91066    0.02736  33.284 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    9.6792e+11
Residual Sum of Squares: 1.8123e+11
R-Squared:      0.81276
Adj. R-Squared: 0.77418
F-statistic: 2665.18 on 2 and 1228 DF, p-value: < 2.22e-16
```

Observation:

The Adjusted R^2 value obtained is good but less when compared with the rest off the models we have build so far. But still this is good enough model.

Random Effect Model:

In this the entities are chosen at random, hence the effect of not including the entity would not be correlated with the dependent variable.

Best Fit Model:

```
plm(MILK ~ COWS + FEED, data = Pooled_Data, model = "random", index =
c("FARM", "YEAR"))
```

Summary:

```
Balanced Panel: n = 247, T = 6, N = 1482

Effects:
                   var   std.dev share
idiosyncratic 152525731     12350 0.392
individual    237041434     15396 0.608
```

```
theta: 0.6888

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-90549.09  -5474.03    263.04   5013.96  67029.95

Coefficients:
               Estimate  Std. Error z-value  Pr(>|z|)
(Intercept) -1.4621e+04  2.0743e+03 -7.0484  1.81e-12 ***
COWS         4.0014e+03  1.2531e+02 31.9315 < 2.2e-16 ***
FEED         9.8777e-01  2.4010e-02 41.1402 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    2.5951e+12
Residual Sum of Squares: 2.2631e+11
R-Squared:      0.91279
Adj. R-Squared: 0.91268
Chisq: 15480.8 on 2 DF, p-value: < 2.22e-16
```

Observation:

Adjusted R^2 value is more than 90% which very high, means that the independent variables are explaining the dependent variable in an efficient manner.

Tests:

Test for Heteroskedasticity:

Breusch-Pagan Test:

Since Entity Fixed (n-1) Binary Variables Model has highest Adjusted R^2 value of 0.9822 we will perform BP Test to see whether Heteroskedasticity exists or not.

The null Hypothesis for BP Test is that the data is Heteroskedasticity.

```
Model:  lm(MILK ~ COWS + FEED + FARM, data=Pooled_Data)
BP = 603.02, df = 248, p-value < 2.2e-16
```

Here, BinFARM_Model3 is the name given to Best Fit Model under Entity Fixed (n-1) Binary Variables.

Since, p-value is less than 0.05, we reject the null hypothesis and establish the fact that there is presence of heteroskedasticity in the data. So, the predicted dependent variables and the residuals of these variables are expected to show heteroskedasticity. To correct this heteroskedasticity-robust standard error can be used.

Test for Serial Collinearity:

Breusch-Godfrey / Wooldridge Test:

The null hypothesis for the Breusch-Godfrey/ Wooldridge test is that there is no serial correlation.

```
data:  MILK ~ COWS + LABOR + FEED
chisq = 521.97, df = 6, p-value < 2.2e-16
alternative hypothesis: serial correlation in idiosyncratic errors
```

Since, p-value is less than 0.05, we reject the null hypothesis and establish the fact that there is presence of serial correlation in the data. This is obvious, as Panel data always has correlation between them.

Test for Panel Effect:

LM Test:

To decide whether the OLS Pooled model is better or the Fixed Effects Model, we use the LM Test. The null hypothesis is that OLS is better than the fixed effects model.

```
data:  MILK ~ COWS + FEED
F = 10.357, df1 = 246, df2 = 1233, p-value < 2.2e-16
alternative hypothesis: significant effects
```

Since p-value is less than 0.05, we reject null hypothesis, thus we can say that the fixed effects model should be chosen ov er the OLS pooling model.

Random Effect and Fixed Effect Model:

Hausman Test:

To decide whether the Fixed effect model is better or the Random effect model, we use the Hausman Test. The null hypothesis is that Fixed Effect model is better than Random Effect model

```
data:  MILK ~ COWS + FEED
chisq = 6.8832, df = 2, p-value = 0.03201
alternative hypothesis: one model is inconsistent
```

Since the p-value is larger than 0.01, we fail to reject the null hypothesis, thus we can say that fixed effect model is better than the random effects model.

Correctness Measure for Heteroskedasticity:

We will be using clustered standard errors & robust standard errors to give more correct values.

This corrects the coefficients.

```
               Estimate  Std. Error  t value  Pr(>|t|)
(Intercept) -1.2438e+04  4.8278e+03  -2.5763 0.0101012 *
COWS         4.1634e+03  3.0208e+02  13.7825 < 2.2e-16 ***
FEED         9.5899e-01  5.5031e-02  17.4263 < 2.2e-16 ***
FARM2        6.5596e+03  4.2563e+03   1.5411 0.1235389
FARM3       -8.9197e+03  3.7671e+03  -2.3678 0.0180472 *
FARM4        1.6687e+04  1.1588e+04   1.4401 0.1500926
FARM5        2.7492e+03  4.5012e+03   0.6108 0.5414646
FARM6        6.6220e+03  3.8120e+03   1.7371 0.0826105 .
FARM7       -5.5481e+03  8.7583e+03  -0.6335 0.5265495
FARM8        8.0586e+02  6.0130e+03   0.1340 0.8934091
FARM9        3.7534e+04  1.7607e+04   2.1318 0.0332206 *
FARM10       2.8589e+04  8.3118e+03   3.4396 0.0006020 ***
FARM11       1.7405e+04  4.5504e+03   3.8250 0.0001373 ***
FARM12       1.6110e+04  5.8457e+03   2.7559 0.0059388 **
FARM13       2.2734e+04  1.1088e+04   2.0504 0.0405371 *
FARM14      -4.0823e+03  6.5612e+03  -0.6222 0.5339381
FARM15      -2.4417e+04  5.7271e+03  -4.2634 2.167e-05 ***
FARM16       6.3531e+03  1.4933e+04   0.4254 0.6705913
FARM17      -1.5758e+04  3.3807e+03  -4.6613 3.484e-06 ***
FARM18      -1.6418e+03  3.2177e+03  -0.5103 0.6099649
FARM19      -1.4843e+04  4.4048e+03  -3.3698 0.0007754 ***
FARM20      -1.9179e+04  5.1251e+03  -3.7422 0.0001908 ***
FARM21       1.1779e+04  4.4410e+03   2.6524 0.0080939 **
FARM22       3.6545e+03  4.4662e+03   0.8183 0.4133721
FARM23      -1.9131e+04  4.5009e+03  -4.2506 2.293e-05 ***
FARM24       5.9605e+03  7.0167e+03   0.8495 0.3957837
FARM25      -6.0512e+03  4.7462e+03  -1.2750 0.2025608
FARM26       8.6636e+03  1.0206e+04   0.8488 0.3961376
FARM27      -2.8005e+04  7.2548e+03  -3.8602 0.0001192 ***
FARM28      -1.2090e+04  3.4485e+03  -3.5057 0.0004717 ***
```

FARM29	-8.2084e+03	5.6439e+03	-1.4544	0.1460938	
FARM30	1.2129e+04	4.3875e+03	2.7644	0.0057873	**
FARM31	4.9445e+04	9.9358e+03	4.9765	7.392e-07	***
FARM32	2.2104e+04	6.8680e+03	3.2184	0.0013228	**
FARM33	3.7078e+04	1.5474e+04	2.3961	0.0167183	*
FARM34	-2.7251e+04	6.6486e+03	-4.0988	4.424e-05	***
FARM35	-1.4105e+04	5.8843e+03	-2.3970	0.0166761	*
FARM36	1.0075e+04	4.1133e+03	2.4495	0.0144456	*
FARM37	-2.4048e+04	6.4615e+03	-3.7218	0.0002067	***
FARM38	-2.3417e+04	1.3152e+04	-1.7804	0.0752506	.
FARM39	-2.3417e+04	5.5327e+03	-4.2324	2.483e-05	***
FARM40	-9.0596e+03	3.6469e+03	-2.4842	0.0131171	*
FARM41	-2.4530e+04	3.8383e+03	-6.3909	2.333e-10	***
FARM42	-1.9694e+04	6.2821e+03	-3.1350	0.0017590	**
FARM43	-1.3417e+04	5.5353e+03	-2.4238	0.0155009	*
FARM44	-1.7538e+04	3.8414e+03	-4.5657	5.477e-06	***
FARM45	-5.1134e+03	5.0142e+03	-1.0198	0.3080293	
FARM46	-2.0587e+04	5.5896e+03	-3.6831	0.0002404	***
FARM47	-6.5601e+03	5.3503e+03	-1.2261	0.2203849	
FARM48	-1.5821e+03	4.8317e+03	-0.3274	0.7433876	
FARM49	-5.8808e+03	4.3105e+03	-1.3643	0.1727287	
FARM50	-5.3479e+04	1.2691e+04	-4.2138	2.694e-05	***
FARM51	-1.6658e+04	3.0166e+03	-5.5221	4.081e-08	***
FARM52	-1.1986e+04	1.3298e+04	-0.9013	0.3675791	
FARM53	-5.2306e+03	7.0550e+03	-0.7414	0.4585955	
FARM54	-2.8719e+04	7.4160e+03	-3.8726	0.0001133	***
FARM55	-1.0576e+04	3.8950e+03	-2.7151	0.0067176	**
FARM56	-2.1344e+04	3.6700e+03	-5.8158	7.679e-09	***
FARM57	6.7842e+03	3.8781e+03	1.7494	0.0804742	.
FARM58	-6.1235e+02	3.9074e+03	-0.1567	0.8754966	
FARM59	-1.0533e+04	3.7560e+03	-2.8044	0.0051197	**
FARM60	-1.1426e+04	6.2679e+03	-1.8229	0.0685555	.
FARM61	8.5847e+03	4.1145e+03	2.0865	0.0371426	*
FARM62	-7.1060e+03	6.6869e+03	-1.0627	0.2881405	
FARM63	3.6625e+04	1.1594e+04	3.1590	0.0016218	**
FARM64	-3.9143e+04	8.1401e+03	-4.8087	1.706e-06	***
FARM65	-5.0773e+03	4.1599e+03	-1.2205	0.2224913	
FARM66	-2.5560e+04	4.9094e+03	-5.2063	2.255e-07	***
FARM67	-7.5997e+03	4.2525e+03	-1.7871	0.0741650	.
FARM68	-1.2058e+04	3.5129e+03	-3.4326	0.0006177	***
FARM69	-7.0177e+04	6.6771e+03	-10.5102	< 2.2e-16	***
FARM70	-2.1481e+04	4.9764e+03	-4.3165	1.713e-05	***
FARM71	5.0241e+03	3.6556e+03	1.3743	0.1695841	
FARM72	-3.0490e+04	5.8870e+03	-5.1792	2.600e-07	***
FARM73	2.4000e+03	6.0334e+03	0.3978	0.6908536	
FARM74	-8.8320e+03	3.3614e+03	-2.6275	0.0087076	**
FARM75	7.2466e+03	8.5261e+03	0.8499	0.3955249	
FARM76	6.4162e+03	4.9698e+03	1.2910	0.1969378	
FARM77	-1.3015e+04	5.4544e+03	-2.3862	0.0171723	*
FARM78	-5.1311e+02	7.0665e+03	-0.0726	0.9421262	
FARM79	-1.9667e+03	3.8029e+03	-0.5171	0.6051460	
FARM80	-7.1902e+03	1.6445e+04	-0.4372	0.6620153	
FARM81	1.4797e+04	1.0886e+04	1.3592	0.1743256	
FARM82	-4.0837e+03	4.3335e+03	-0.9424	0.3461890	
FARM83	-1.7024e+04	4.3547e+03	-3.9093	9.760e-05	***
FARM84	-8.0587e+02	4.1245e+03	-0.1954	0.8451217	
FARM85	-2.3111e+04	2.4834e+04	-0.9306	0.3522353	
FARM86	1.0574e+04	9.8756e+03	1.0707	0.2845018	
FARM87	3.5397e+04	9.6567e+03	3.6655	0.0002573	***
FARM88	2.9480e+04	1.0954e+04	2.6912	0.0072155	**
FARM89	-3.3163e+03	5.7325e+03	-0.5785	0.5630237	
FARM90	4.4165e+03	4.2322e+03	1.0435	0.2969023	
FARM91	-6.3301e+03	1.2858e+04	-0.4923	0.6225962	
FARM92	-6.5073e+03	7.9288e+03	-0.8207	0.4119668	
FARM93	-2.8666e+03	9.0523e+03	-0.3167	0.7515437	

FARM94	1.6966e+04	9.5913e+03	1.7689	0.0771575	.
FARM95	-6.1932e+03	4.9584e+03	-1.2490	0.2118928	
FARM96	7.7638e+03	1.0046e+04	0.7729	0.4397537	
FARM97	1.1002e+03	5.8067e+03	0.1895	0.8497539	
FARM98	-4.3539e+02	4.1472e+03	-0.1050	0.9164052	
FARM99	-9.9520e+03	6.2529e+03	-1.5916	0.1117356	
FARM100	6.5721e+03	3.9466e+03	1.6652	0.0961201	.
FARM101	1.8818e+04	3.5271e+03	5.3354	1.133e-07	***
FARM102	-1.1912e+04	7.7866e+03	-1.5298	0.1263268	
FARM103	-3.5379e+03	1.9409e+04	-0.1823	0.8553929	
FARM104	-1.2981e+04	5.8783e+03	-2.2083	0.0274086	*
FARM105	-1.8148e+04	3.7798e+03	-4.8014	1.769e-06	***
FARM106	-4.7742e+03	4.3912e+03	-1.0872	0.2771532	
FARM107	-4.9527e+03	8.3258e+03	-0.5949	0.5520413	
FARM108	1.4400e+04	8.0534e+03	1.7881	0.0740132	.
FARM109	-4.9719e+04	1.1611e+04	-4.2822	1.994e-05	***
FARM110	6.4337e+03	5.2091e+03	1.2351	0.2170356	
FARM111	-9.4789e+03	6.2273e+03	-1.5221	0.1282318	
FARM112	-7.4162e+03	3.3395e+03	-2.2208	0.0265490	*
FARM113	7.4817e+03	4.0466e+03	1.8489	0.0647118	.
FARM114	6.3238e+03	3.6789e+03	1.7190	0.0858737	.
FARM115	6.9230e+03	3.8028e+03	1.8205	0.0689244	.
FARM116	-4.4041e+03	6.9389e+03	-0.6347	0.5257439	
FARM117	-5.6240e+03	3.3373e+03	-1.6852	0.0922009	.
FARM118	-7.0029e+03	4.2007e+03	-1.6671	0.0957483	.
FARM119	1.4226e+03	3.6584e+03	0.3888	0.6974544	
FARM120	-5.1002e+03	3.5751e+03	-1.4266	0.1539543	
FARM121	1.3269e+04	6.1080e+03	2.1724	0.0300145	*
FARM122	3.4525e+03	3.6041e+03	0.9579	0.3382854	
FARM123	-2.0340e+03	3.6268e+03	-0.5608	0.5750077	
FARM124	-1.9952e+03	4.1080e+03	-0.4857	0.6272767	
FARM125	5.4431e+03	3.8157e+03	1.4265	0.1539749	
FARM126	-1.1478e+04	6.4883e+03	-1.7690	0.0771359	.
FARM127	-1.7361e+03	3.2783e+03	-0.5296	0.5965085	
FARM128	-1.3433e+04	5.0848e+03	-2.6419	0.0083498	**
FARM129	1.4446e+04	4.9505e+03	2.9180	0.0035867	**
FARM130	-2.0043e+03	3.3911e+03	-0.5910	0.5546020	
FARM131	5.0754e+03	4.9427e+03	1.0268	0.3046926	
FARM132	6.9686e+03	3.9615e+03	1.7591	0.0788086	.
FARM133	-1.1177e+04	4.7757e+03	-2.3403	0.0194248	*
FARM134	2.3799e+03	5.0818e+03	0.4683	0.6396314	
FARM135	6.5725e+03	4.7252e+03	1.3909	0.1644944	
FARM136	1.1937e+04	4.9791e+03	2.3975	0.0166569	*
FARM137	-1.0524e+04	8.6111e+03	-1.2221	0.2219004	
FARM138	-2.6136e+03	3.6942e+03	-0.7075	0.4793915	
FARM139	1.0081e+04	4.6538e+03	2.1662	0.0304849	*
FARM140	-7.3502e+03	4.5126e+03	-1.6288	0.1036067	
FARM141	-7.3000e+03	3.7118e+03	-1.9667	0.0494405	*
FARM142	-8.1647e+03	8.2597e+03	-0.9885	0.3231028	
FARM143	2.7285e+03	4.3446e+03	0.6280	0.5301144	
FARM144	5.1001e+03	3.4071e+03	1.4969	0.1346756	
FARM145	-1.9301e+04	6.8204e+03	-2.8298	0.0047326	**
FARM146	8.4639e+03	3.8041e+03	2.2250	0.0262649	*
FARM147	-8.9324e+02	3.7130e+03	-0.2406	0.8099285	
FARM148	-3.9657e+03	3.6277e+03	-1.0932	0.2745333	
FARM149	1.5435e+04	3.3767e+03	4.5710	5.343e-06	***
FARM150	6.2325e+03	3.8080e+03	1.6367	0.1019473	
FARM151	2.0770e+03	4.2920e+03	0.4839	0.6285266	
FARM152	-7.0800e+03	3.4835e+03	-2.0325	0.0423208	*
FARM153	-1.8048e+04	5.0756e+03	-3.5559	0.0003910	***
FARM154	-3.9930e+03	7.5627e+03	-0.5280	0.5976059	
FARM155	7.8176e+03	4.8015e+03	1.6282	0.1037422	
FARM156	1.0585e+04	1.6310e+04	0.6490	0.5164810	
FARM157	-4.8421e+04	6.3675e+03	-7.6043	5.657e-14	***
FARM158	9.3166e+03	6.7323e+03	1.3839	0.1666490	

```
FARM159     -8.9309e+03  4.9469e+03  -1.8054 0.0712610 .
FARM160     -3.3060e+04  6.8784e+03  -4.8063 1.727e-06 ***
FARM161     -1.3606e+03  4.9611e+03  -0.2743 0.7839334
FARM162      3.3847e+03  3.6928e+03   0.9166 0.3595501
FARM163     -1.2661e+04  4.4147e+03  -2.8679 0.0042025 **
FARM164     -5.3297e+04  1.3689e+04  -3.8934 0.0001042 ***
FARM165     -9.1711e+03  6.3405e+03  -1.4464 0.1483108
FARM166      3.8129e+04  9.5190e+03   4.0056 6.556e-05 ***
FARM167     -3.5511e+04  5.5239e+03  -6.4285 1.837e-10 ***
FARM168      9.4738e+02  3.8196e+03   0.2480 0.8041513
FARM169     -8.4684e+03  3.9126e+03  -2.1644 0.0306269 *
FARM170     -3.1665e+03  3.6091e+03  -0.8774 0.3804613
FARM171     -1.2621e+04  3.3881e+03  -3.7252 0.0002040 ***
FARM172     -8.9052e+03  4.0847e+03  -2.1801 0.0294360 *
FARM173     -9.3143e+03  3.6906e+03  -2.5238 0.0117340 *
FARM174      7.0135e+03  3.7082e+03   1.8914 0.0588086 .
FARM175      2.9989e+03  3.9181e+03   0.7654 0.4441890
FARM176     -1.6653e+04  5.6167e+03  -2.9650 0.0030855 **
FARM177      2.0975e+04  7.0133e+03   2.9908 0.0028376 **
FARM178      5.2953e+03  3.4171e+03   1.5496 0.1214885
FARM179      2.8853e+03  3.9003e+03   0.7398 0.4595867
FARM180      2.0885e+03  5.2797e+03   0.3956 0.6924915
FARM181     -1.9563e+04  4.8648e+03  -4.0214 6.137e-05 ***
FARM182      7.8298e+03  3.8380e+03   2.0401 0.0415580 *
FARM183     -8.6424e+03  7.1902e+03  -1.2020 0.2296095
FARM184     -5.3490e+04  1.2943e+04  -4.1328 3.825e-05 ***
FARM185     -8.7631e+02  4.1609e+03  -0.2106 0.8332279
FARM186     -8.8723e+03  9.1948e+03  -0.9649 0.3347711
FARM187     -4.8046e+03  4.5780e+03  -1.0495 0.2941581
FARM188     -9.7767e+03  6.8608e+03  -1.4250 0.1544074
FARM189      4.5854e+03  5.0584e+03   0.9065 0.3648506
FARM190      6.6734e+03  3.3484e+03   1.9930 0.0464834 *
FARM191      1.2323e+04  3.8845e+03   3.1724 0.0015491 **
FARM192     -1.8311e+02  4.0247e+03  -0.0455 0.9637188
FARM193      1.2088e+04  5.3737e+03   2.2495 0.0246592 *
FARM194     -6.5317e+03  3.5698e+03  -1.8297 0.0675327 .
FARM195     -1.2725e+04  3.9430e+03  -3.2272 0.0012828 **
FARM196     -8.0374e+03  1.3318e+04  -0.6035 0.5462840
FARM197      2.3072e+03  3.9327e+03   0.5867 0.5575419
FARM198     -3.6370e+03  3.4932e+03  -1.0412 0.2980060
 [ reached getOption("max.print") -- omitted 49 rows ]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Conclusion:

- On conducting Hausman Test, we confirm that Fixed Effect Models perform better than Random Effect Model at 99% significance level, but it is not true when reduce the significance level to 95%, this is evident as Random effects model presumes that there is no fixed effects & that all variations are caused due to random effects, which clearly is not the case here as each FARM has its own entity specific effects.
- Among all the models develop, we can say that Entity Fixed with n-1 Binary Variables is the best, as it has the highest Adjusted R^2 value, this means that this model is able to explain dependent Variable most efficiently.
- Also, from LM test we can say that Fixed Effect Models are performing better than OLS Pooling Model, this is also evident when we compare the Adjusted R^2 value two Models.
- The data is Heteroskedastic as observed from BP test, and also there exists Serial Collinearity Between Variables, this issue is handled by using clustered standard errors & robust standard errors to give more correct values.

Reference:

- Dataset : http://people.stern.nyu.edu/wgreene/Econometrics/dairy.csv

- https://www.rdocumentation.org/packages/plm/versions/1.6-5/topics/phtest
- https://en.wikipedia.org/wiki/Panel_analysis

R code:

```
# Loading Libraries
library(ggplot2)
library(olsrr)
library(gvlma)
library(car)
library(gplots)
library(plm)
library(sandwich)
library(lmtest)
library(corrplot)

data <- read.csv("E:\\IIT Kanpur\\IME 2nd Sem Courses\\MBA652-SMBA\\Projects\\Panel Data Regression\\Spanish Dairy\\dairy_1.csv")

data_1 <- data
data_1$FARM <- NULL
data_1$YEAR <- NULL

# Finding Correlation Between varibales
correlation <- cor(data_1)
correlation_1 <- round(correlation, 2)
corrplot(correlation_1, method = "number")

data$YIT <- NULL

summary(data)
sd(data$MILK)
help(sd)

sum(is.na("data"))


# Scatter Plots
ggplot(data, aes(x = COWS, y = MILK)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Milk Produced vs Cows in the FARM", x = "COWS", y = "MILK")

ggplot(data, aes(x = LAND, y = MILK)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Milk Produced vs Land of the FARM", x = "LAND AREA", y = "MILK")

ggplot(data, aes(x = LABOR, y = MILK)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Milk Produced vs Labor Involved", x = "Labor", y = "MILK")

ggplot(data, aes(x = FEED, y = MILK)) +
  geom_point() +
  geom_smooth(method = "lm") +
```

```
  labs(title = "Milk Produced vs Quantity of food fed to Cows", x = "Quantitiy Fed", y = "MILK")



ggplot(data= data, aes(x = FARM, y = MILK, group = YEAR, colour = YEAR))+
  geom_point() +
  geom_smooth(method= "lm", se= TRUE) +
  labs(x= "FARM Number", y= "Amount of Milk Produced", title = "Milk Production vs Farm")


ggplot(data= data, aes(x = YEAR, y = MILK, group = FARM, colour = FARM))+
  geom_point() +
  geom_smooth(method= "lm", se= TRUE) +
  labs(x= "Year", y= "Amount of Milk Produced", title = "Milk Production vs Year")

#deviation over years
plotmeans(MILK ~ YEAR, error.bars="sd", main="Deviations-Years", data=data, xlab="YEAR", ylab="Milk
Production" )

#deviation over country
plotmeans(MILK ~ FARM, main="Deviations-FARMs", data=data, xlab="FARM", ylab= "Milk Production" )

#Regression Models
# Pooling Model
Pooled_Data <- pdata.frame(data, index = c("FARM","YEAR"))

Pooled_model1 <- plm(MILK ~ COWS + LAND + LABOR + FEED, data=Pooled_Data, model = 'pooling')
summary(Pooled_model1)
bptest(Pooled_model1)

Pooled_model2 <- plm(MILK ~ COWS + FEED + LABOR, data=Pooled_Data, model = 'pooling')
summary(Pooled_model2)
bptest(Pooled_model2)

#best model
Pooled_model3 <- plm(MILK ~ COWS + FEED, data=Pooled_Data, model = 'pooling')
summary(Pooled_model3)
bptest(Pooled_model3)



# Entity Fixed Models
# 1. Entity Demeaned Models
EF_Model1 <- plm(MILK ~ COWS + LAND + LABOR + FEED, data=Pooled_Data, model = 'within', effect =
'individual')
summary(EF_Model1)

EF_Model2 <- plm(MILK ~ COWS + LABOR + FEED, data=Pooled_Data, model = 'within', effect = 'individual')
summary(EF_Model2)

#best model
EF_Model3 <- plm(MILK ~ COWS + FEED, data=Pooled_Data, model = 'within', effect = 'individual')
summary(EF_Model3)
```

```
# 2. Binary Variables for FARM
BinFARM_Model1 <- lm(MILK ~ COWS + LAND + LABOR + FEED + FARM, data=Pooled_Data)
summary(BinFARM_Model1)

BinFARM_Model2 <- lm(MILK ~ COWS + LABOR + FEED + FARM, data=Pooled_Data)
summary(BinFARM_Model1)

# Best Model
BinFARM_Model3 <- lm(MILK ~ COWS + FEED + FARM, data=Pooled_Data)
summary(BinFARM_Model1)



# Time Fixed Models
# 1. Binary Variables for year
BinYEAR_Model1 <- lm(MILK ~ COWS + LAND + LABOR + FEED + YEAR, data=Pooled_Data)
summary(BinYEAR_Model1)

# best model
BinYEAR_Model2 <- lm(MILK ~ COWS + LABOR + FEED + YEAR, data=Pooled_Data)
summary(BinYEAR_Model2)



# 2. Time Demeaned Models
TF_Model1 <- plm(MILK ~ COWS + LAND + LABOR + FEED, data=Pooled_Data, model = 'within', effect = 'time')
summary(TF_Model1)

#best Model
TF_Model2 <- plm(MILK ~ COWS + LABOR + FEED, data=Pooled_Data, model = 'within', effect = 'time')
summary(TF_Model2)



# Entity and Time Fixed Models
ETF_Model1 <- plm(MILK ~ COWS + LAND + LABOR + FEED, data=Pooled_Data, model =
'within',index=c("FARM","YEAR"), effect ='twoways')
summary(ETF_Model1)

ETF_Model2 <- plm(MILK ~ COWS + LABOR + FEED, data=Pooled_Data, model =
'within',index=c("FARM","YEAR"), effect ='twoways')
summary(ETF_Model2)

#Best Model
ETF_Model3 <- plm(MILK ~ COWS + FEED, data=Pooled_Data, model = 'within',index=c("FARM","YEAR"), effect
='twoways')
summary(ETF_Model3)



#First Difference Model
```

```
# best Model
FD_Model1 <- plm(MILK ~ COWS + LAND + LABOR + FEED, data=Pooled_Data, index=c("FARM","YEAR"),model
= 'fd')
summary(FD_Model1)

FD_Model2 <- plm(MILK ~ COWS + LAND + FEED, data=Pooled_Data, index=c("FARM","YEAR"),model = 'fd')
summary(FD_Model2)

FD_Model3 <- plm(MILK ~ COWS + FEED, data=Pooled_Data, index=c("FARM","YEAR"),model = 'fd')
summary(FD_Model3)
#None of the above models can be used




# Random Effect Model
RF_Model1 <- plm(MILK ~ COWS + LAND + LABOR + FEED, data = Pooled_Data, model = "random", index =
c("FARM", "YEAR"))
summary(RF_Model1)

RF_Model2 <- plm(MILK ~ COWS + LABOR + FEED, data = Pooled_Data, model = "random", index = c("FARM",
"YEAR"))
summary(RF_Model2)

# best Model
RF_Model3 <- plm(MILK ~ COWS + FEED, data = Pooled_Data, model = "random", index = c("FARM", "YEAR"))
summary(RF_Model3)




#Chosen Models
# Pooling Model
#best model
Pooled_model3 <- plm(MILK ~ COWS + FEED, data=Pooled_Data, model = 'pooling')
summary(Pooled_model3)
bptest(Pooled_model3)

# Entity Fixed Models
# Entity Demeaned Model
#best model
EF_Model3 <- plm(MILK ~ COWS + FEED, data=Pooled_Data, model = 'within', effect = 'individual')
summary(EF_Model3)

# Binary variables for FARM
# Best Model
BinFARM_Model3 <- lm(MILK ~ COWS + FEED + FARM, data=Pooled_Data)
summary(BinFARM_Model3)
bptest(BinFARM_Model3)
coeftest(BinFARM_Model3, vcovHC(BinFARM_Model3, method = "arellano"))

# Time Fixed Models
# Binary Variables for YEAR
# best model
BinYEAR_Model2 <- lm(MILK ~ COWS + LABOR + FEED + YEAR, data=Pooled_Data)
```

```
summary(BinYEAR_Model2)

# Time Demeaned Model
#best Model
TF_Model2 <- plm(MILK ~ COWS + LABOR + FEED, data=Pooled_Data, model = 'within', effect = 'time')
summary(TF_Model2)
pFtest(TF_Model2, Pooled_model3)

pbgtest(TF_Model2)
plmtest(TF_Model2,c("time"), type=("bp"))

# Entity and Time Fixed Models
# Best Model
ETF_Model3 <- plm(MILK ~ COWS + FEED, data=Pooled_Data, model = 'within',index=c("FARM","YEAR"), effect
='twoways')
summary(ETF_Model3)

# Random Effect Model
# best Model
RF_Model3 <- plm(MILK ~ COWS + FEED, data = Pooled_Data, model = "random", index = c("FARM", "YEAR"))
summary(RF_Model3)



# Tests
# BP test heteroskedasticity
bptest(BinFARM_Model3)

# Breusch-Godfrey / Wooldridge Test Serail Collinearity
pbgtest(TF_Model2)

# LM TEST test for Panel Effect:
pFtest(EF_Model3, Pooled_model3)

# Hausman test
phtest(RF_Model3, BinFARM_Model3)

# Heteroskedasticity correctness measures
coeftest(BinFARM_Model3, vcovHC(BinFARM_Model3, method = "arellano"))
```