\*The Six Cities Study of Air Pollution and Health was a longitudinal study designed to characterize lung growth as measured by changes in pulmonary function in children and adolescents, and the factors that influence lung function growth.

A cohort of 13,379 children born on or after 1967 was enrolled in six communities across the U.S.: Watertown (Massachusetts), Kingston and Harriman (Tennessee), a section of St. Louis (Missouri), Steubenville (Ohio), Portage (Wisconsin), and Topeka (Kansas). Most children were enrolled in the first or second grade (between the ages of six and seven) and measurements of study participants were obtained annually until graduation from high school or loss to follow-up. At each annual examination, spirometry, the measurement of pulmonary function, was performed and a respiratory health questionnaire was completed by a parent or guardian.

**data** air\_pol;

input ID Height Age INI\_Height INI\_Age Log\_FEV1;

datalines;

1 1.20 9.3415 1.20 9.3415 0.21511

…………………………….

300 1.63 17.8645 1.44 11.9617 1.16938

;

**run**;

\*First, we'll take a couple of snapshots of the data;

ods rtf file="E:\12 - Case Study.rtf";

**Proc** **SGplot** data = air\_pol;

series x=Age y=Log\_FEV1 / group =ID LineAttrs= (pattern=**1** );

**run**;



**Proc** **SGplot** data = air\_pol;

series x=Height y=Log\_FEV1 / group =ID LineAttrs= (pattern=**1** );

**run**;



**data** air\_pol\_trans;

set air\_pol;

L\_Height = log(Height);

L\_INI\_Height = log(INI\_Height);

L\_Age = log(Age);

L\_INI\_Age = log(INI\_Age);

**run**;

**proc** **mixed** data = air\_pol\_trans;

class ID;

model Log\_FEV1 = INI\_Height Height INI\_Age Age/ solution outp=air\_pol\_age\_pred;

random intercept Age/type=UN subject=ID g gcorr v vcorr;

**run**;

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | -4633.5 |
| **AIC (Smaller is Better)** | -4625.5 |
| **AICC (Smaller is Better)** | -4625.5 |
| **BIC (Smaller is Better)** | -4610.7 |

**proc** **mixed** data = air\_pol\_trans;

class ID;

model Log\_FEV1 = INI\_Height Height INI\_Age Age/ solution outp=air\_pol\_age\_pred;

random intercept Height/type=UN subject=ID g gcorr v vcorr s;

ods output SolutionR=RE\_est;

**run**;

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | -4653.6 |
| **AIC (Smaller is Better)** | -4645.6 |
| **AICC (Smaller is Better)** | -4645.5 |
| **BIC (Smaller is Better)** | -4630.8 |

**proc** **mixed** data = air\_pol\_trans;

class ID;

model Log\_FEV1 = INI\_Height Height L\_INI\_Age L\_Age/ solution;

random intercept Height/type=UN subject=ID g gcorr v vcorr;

**run**;

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | -4642.6 |
| **AIC (Smaller is Better)** | -4634.6 |
| **AICC (Smaller is Better)** | -4634.6 |
| **BIC (Smaller is Better)** | -4619.8 |

proc mixed data = air\_pol\_trans;

class ID;

model Log\_FEV1 = INI\_Height Height L\_INI\_Age L\_Age/ solution outp=air\_pol\_height\_pred;

random intercept Height/type=UN subject=ID g gcorr v vcorr;

run;

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | -4648.1 |
| **AIC (Smaller is Better)** | -4640.1 |
| **AICC (Smaller is Better)** | -4640.0 |
| **BIC (Smaller is Better)** | -4625.3 |

**proc** **mixed** data = air\_pol\_trans;

class ID;

model Log\_FEV1 = L\_INI\_Height L\_Height INI\_Age Age/ solution outp=air\_pol\_age\_pred;

random intercept Age/type=UN subject=ID g gcorr v vcorr;

**run**;

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | -4567.9 |
| **AIC (Smaller is Better)** | -4559.9 |
| **AICC (Smaller is Better)** | -4559.9 |
| **BIC (Smaller is Better)** | -4545.1 |

**proc** **mixed** data = air\_pol\_trans;

class ID;

model Log\_FEV1 = L\_INI\_Height L\_Height INI\_Age Age/ solution;

random intercept L\_Height/type=UN subject=ID g gcorr v vcorr;

**run**;

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | -4589.5 |
| **AIC (Smaller is Better)** | -4581.5 |
| **AICC (Smaller is Better)** | -4581.5 |
| **BIC (Smaller is Better)** | -4566.7 |

|  |
| --- |
| The SAS System |

The Mixed Procedure

| **Model Information** | |
| --- | --- |
| **Data Set** | WORK.AIR\_POL\_TRANS |
| **Dependent Variable** | Log\_FEV1 |
| **Covariance Structure** | Unstructured |
| **Subject Effect** | ID |
| **Estimation Method** | REML |
| **Residual Variance Method** | Profile |
| **Fixed Effects SE Method** | Model-Based |
| **Degrees of Freedom Method** | Containment |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **ID** | 299 | 1 2 3 …………… 300 |

| **Dimensions** | |
| --- | --- |
| **Covariance Parameters** | 4 |
| **Columns in X** | 5 |
| **Columns in Z per Subject** | 2 |
| **Subjects** | 299 |
| **Max Obs per Subject** | 12 |

| **Number of Observations** | |
| --- | --- |
| **Number of Observations Read** | 1993 |
| **Number of Observations Used** | 1993 |
| **Number of Observations Not Used** | 0 |

| **Iteration History** | | | |
| --- | --- | --- | --- |
| **Iteration** | **Evaluations** | **-2 Res Log Like** | **Criterion** |
| **0** | 1 | -2969.35809510 |  |
| **1** | 2 | -4653.34236076 | 0.00005125 |
| **2** | 1 | -4653.56627039 | 0.00000036 |
| **3** | 1 | -4653.56775349 | 0.00000000 |

|  |
| --- |
| Convergence criteria met. |

| **Estimated G Matrix** | | | | |
| --- | --- | --- | --- | --- |
| **Row** | **Effect** | **ID** | **Col1** | **Col2** |
| **1** | Intercept | 1 | 0.06835 | -0.04607 |
| **2** | Height | 1 | -0.04607 | 0.03523 |

| **Estimated G Correlation Matrix** | | | | |
| --- | --- | --- | --- | --- |
| **Row** | **Effect** | **ID** | **Col1** | **Col2** |
| **1** | Intercept | 1 | 1.0000 | -0.9388 |
| **2** | Height | 1 | -0.9388 | 1.0000 |

| **Estimated V Matrix for ID 1** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Col1** | **Col2** | **Col3** | **Col4** | **Col5** | **Col6** | **Col7** |
| **1** | 0.01190 | 0.008217 | 0.008027 | 0.007686 | 0.007459 | 0.007383 | 0.007307 |
| **2** | 0.008217 | 0.01152 | 0.008090 | 0.008003 | 0.007945 | 0.007925 | 0.007906 |
| **3** | 0.008027 | 0.008090 | 0.01151 | 0.008201 | 0.008248 | 0.008264 | 0.008280 |
| **4** | 0.007686 | 0.008003 | 0.008201 | 0.01194 | 0.008795 | 0.008874 | 0.008953 |
| **5** | 0.007459 | 0.007945 | 0.008248 | 0.008795 | 0.01254 | 0.009281 | 0.009402 |
| **6** | 0.007383 | 0.007925 | 0.008264 | 0.008874 | 0.009281 | 0.01280 | 0.009552 |
| **7** | 0.007307 | 0.007906 | 0.008280 | 0.008953 | 0.009402 | 0.009552 | 0.01308 |

| **Estimated V Correlation Matrix for ID 1** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Col1** | **Col2** | **Col3** | **Col4** | **Col5** | **Col6** | **Col7** |
| **1** | 1.0000 | 0.7018 | 0.6858 | 0.6448 | 0.6105 | 0.5982 | 0.5856 |
| **2** | 0.7018 | 1.0000 | 0.7026 | 0.6824 | 0.6610 | 0.6527 | 0.6440 |
| **3** | 0.6858 | 0.7026 | 1.0000 | 0.6996 | 0.6865 | 0.6809 | 0.6747 |
| **4** | 0.6448 | 0.6824 | 0.6996 | 1.0000 | 0.7188 | 0.7179 | 0.7164 |
| **5** | 0.6105 | 0.6610 | 0.6865 | 0.7188 | 1.0000 | 0.7326 | 0.7341 |
| **6** | 0.5982 | 0.6527 | 0.6809 | 0.7179 | 0.7326 | 1.0000 | 0.7382 |
| **7** | 0.5856 | 0.6440 | 0.6747 | 0.7164 | 0.7341 | 0.7382 | 1.0000 |

| **Covariance Parameter Estimates** | | |
| --- | --- | --- |
| **Cov Parm** | **Subject** | **Estimate** |
| **UN(1,1)** | **ID** | 0.06835 |
| **UN(2,1)** | **ID** | -0.04607 |
| **UN(2,2)** | **ID** | 0.03523 |
| **Residual** |  | 0.003381 |

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | -4653.6 |
| **AIC (Smaller is Better)** | -4645.6 |
| **AICC (Smaller is Better)** | -4645.5 |
| **BIC (Smaller is Better)** | -4630.8 |

| **Null Model Likelihood Ratio Test** | | |
| --- | --- | --- |
| **DF** | **Chi-Square** | **Pr > ChiSq** |
| 3 | 1684.21 | <.0001 |

| **Solution for Fixed Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Effect** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | -1.9495 | 0.09604 | 296 | -20.30 | <.0001 |
| **INI\_Height** | 0.1449 | 0.1112 | 1441 | 1.30 | 0.1926 |
| **Height** | 1.6424 | 0.03268 | 251 | 50.26 | <.0001 |
| **INI\_Age** | -0.01478 | 0.007511 | 1441 | -1.97 | 0.0493 |
| **Age** | 0.01929 | 0.001279 | 1441 | 15.08 | <.0001 |

| **Solution for Random Effects** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Effect** | **ID** | **Estimate** | **Std Err Pred** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | **1** | -0.1015 | 0.1875 | 1441 | -0.54 | 0.5882 |
| **Height** | **1** | 0.08522 | 0.1342 | 1441 | 0.63 | 0.5256 |
| **Intercept** | **2** | 0.3719 | 0.1476 | 1441 | 2.52 | 0.0119 |
| **Height** | **2** | -0.2139 | 0.1015 | 1441 | -2.11 | 0.0353 |
| **Intercept** | **3** | 0.5153 | 0.1461 | 1441 | 3.53 | 0.0004 |
| **Height** | **3** | -0.2864 | 0.1016 | 1441 | -2.82 | 0.0049 |
| **Intercept** | **300** | -0.08366 | 0.2307 | 1441 | -0.36 | 0.7170 |
| **Height** | **300** | 0.07121 | 0.1482 | 1441 | 0.48 | 0.6310 |

| **Type 3 Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **INI\_Height** | 1 | 1441 | 1.70 | 0.1926 |
| **Height** | 1 | 251 | 2526.46 | <.0001 |
| **INI\_Age** | 1 | 1441 | 3.87 | 0.0493 |
| **Age** | 1 | 1441 | 227.52 | <.0001 |

**Proc** **SGplot** data = air\_pol\_height\_pred;

series x=Age y=pred / group =ID LineAttrs= (pattern=**1** );

**run**;

