Analyzing Production of an Automotive Adhesive

A Solidify Understanding Task

Purpose: *To determine change over time of an adhesive, investigate potential causes, and develop a potential solution.*

Career Field: Data Analytics

Company: Parker-LORD

WTCC Associate Program of Study and Contact Person
Business Analytics
Norene Kemp nckemp@waketech.edu

NC Math 4 Standards
AF.5.1 Construct regression models of linear, quadratic, exponential, logarithmic, & sinusoidal functions of bivariate data using technology to model data and solve problems.
AF.5.2 Compare residuals and residual plots of non-linear models to assess the goodness-of-fit of the model.

Unit Alignment
NC Math 4 Unit 2 – Functions
NC Math 4 Unit 5 – Exploratory Data Analysis
WTCC Math 152 Post-Unit 3 – Correlation and Regression

Common Core State Standards for Mathematical Practice

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.
4. Model with mathematics.
5. Use appropriate tools strategically.
6. Attend to precision.

Prerequisite Skills

- Correlation
- Linear regression
- Excel formula basics and autofill (or equivalent for statistical tool of choice)

Time Required
The time required to complete the activity is approximately 150 minutes, split over two class meetings.
If necessary, Part 3 may be omitted to fit a 90 minute class meeting.

Materials Needed

- Student Activity Sheet
- Data file (provided)
- Microsoft Excel, other statistical software, or calculator

The Teaching Cycle

Desmos Activity (15 minutes): **https://teacher.desmos.com/activitybuilder/custom/610976f8946890717dbce110**

Video Link (also embedded in Desmos activity): **https://youtu.be/TBDoyTS1Hq4**

Launch: Students should work in groups of 2-3. First ask students to complete the Desmos launch activity. This activity reviews previous content that will be used in the activity. The final slide is a link to the launch video introducing students to the activity.

Instruct students to open the provided data file in Microsoft Excel (or available software/calculator). Students should review the variable definitions in the Student Activity Sheet if they are unsure what any of the data columns represent. Students may not be familiar with work shifts (day, late evening, early morning) or dew point (temperature below which water may condense into dew droplets).

Part 1 – Confirming the Problem (15 minutes)

Explore: Students work in their groups to answer each question. This should be a relatively quick introduction to the data set and the problem. Students should find that there is an increase in PX over time and that over half of the batches would be rejected.

Discuss: Either as a whole class or by visiting individual groups, discuss students' observations with them.

Problem 1: Students could visually inspect or subtract PX Recheck – PX Initial to look for positive/negative values.

Problem 2: Sorting the recheck PX column is a quick way to check this. Students don't need the 150 PSI lower limit since we're only concerned about the upper limit. Since Excel counts the headings as row 1, students may accidentally count 21 total data points instead of 20. One observation (209) is close to the spec limit of 210, so students may ask whether they should include it. (Technically not but they all may increase more over time!) Students might provide a quantity of rejected batches instead of a percentage, which can still help them conceptualize the problem at hand.

Problem 3: This is very open-ended; students may hypothesize all sorts of relationships in the variables. Focus on correct calculation of correlation coefficient and interpretation (strong/weak positive/negative relationships).

Part 2 – Modeling PX Drift (50 minutes)

Explore: Students work in their groups to answer each question. This is likely the most time-consuming part of the activity. Students will develop a linear regression, predicted values and residuals. There appears to be a relatively strong relationship between total percent water and drift per day. Once confirming that the linear relationship is appropriate, students compute shelf life (using two different spec limits) for each data entry in the sample.

Discuss: Either as a whole class or by visiting individual groups, discuss students' observations with them.

IMPORTANT: Early in this part, be sure that groups are making calculations efficiently (autofill formulas instead of computing values for each batch by hand). This will save a lot of time throughout the activity!

Problem 4: Students may not be sure which variables to use and/or how to combine them. Check that students understand how their subtraction order affects the sign of the resulting data (recheck – initial is convention, in which case a positive value indicates a PX increase, and vice versa).

Problem 5: Details about "correct formatting" may vary by instructor. Be sure students realize the difference between r^2 (reported by Excel and some other tools) and r. This problem has a strong correlation, particularly in this industry example. Reinforce understanding by asking what each variable (column in the data set) in the regression equation represents. This will help for the upcoming problems.

Pay attention to the technology tools used to generate the regression line, as different tools may round the slope and intercept of the regression line differently. With such small %H2O values, a small change in precision could cause a sizable difference in predicted values.

Problem 6: Check for subtraction order (actual value – predicted value). No noticeable pattern in the residuals suggests that the linear relationship is appropriate.

Part 3 – Applying the Linear Drift Model (60 minutes)

Explore: Students work in their groups to answer each question. Students will use data analysis to make a decision for the manufacturing company. Students will also compare observed PX values of a new batch to those predicted by the regression line. Work from the previous parts culminates here, where students assess the limitations of the linear model and what further data collection and analysis could be done.

Discuss: Either as a whole class or by visiting individual groups, discuss students' observations with them.

Problems 7-9:
These two are tough questions! Students may have trouble calculating shelf life – you can scaffold this as you feel necessary for your class. The end goal would be the formula:

Shelf life in days = ((upper spec limit – initial PX)/drift per day)

To help students develop this formula, you could encourage students to choose one batch to investigate first. How far away is the initial PX from the upper limit (upper spec limit – initial PX)? How many days will it take to get there (divide by drift/day)? Tracking units might help here, too: PSI / (PSI/days) -> days. Rounding here could be to the nearest day or tenth of a day. Using actual vs predicted drift per day values will yield slightly different results. The Parker-LORD representative used the actual data here.

Problem 10: When entering their own %H2O values, students may accidentally misplace decimals, e.g. 3% instead of 0.03%. The reasoning behind students' decisions would be an interesting discussion point.

Problem 11: Students should realize that their linear model is not expected to be accurate when extrapolated outside of the 20-60 day window. Collecting data with longer-term rechecks and analyzing further may provide more useful conclusions. (In reality, Parker-LORD found that the rate of drift tended to decrease over time, meaning that logarithmic and similar models were more appropriate for drift beyond 60 days.)

Part 4 -- Extension (10 minutes)

Explore: Students work in their groups to answer each question. This section allows students to use their knowledge in a similar situation (same setting but analyzing a new batch).

Discuss: Either as a whole class or by visiting individual groups, discuss students' observations with them.

Problem 12: This one new batch that was rechecked four times, not four new batches. Students should calculate drift per day for each recheck (they will be different). Parker-LORD found that the eventual decreasing trend in drift per day continued beyond this time frame (more on that in the last problem).

Problem 13: By using a linear regression we assume that the drift per day is the same regardless of when the recheck occurs.