(The attached PDF file has better formatting.)

Some candidates wonder what they can model with ARIMA processes. The NEAS web site has project templates for financial series (interest rates, CPI, GDP, unemployment, gas prices), daily temperature, sports won-loss records, and other topics

These project templates have suggestions for time series analysis and ARIMA fitting. You have hundreds of time series to choose from. You can take ideas from the NEAS postings or past student projects and apply them to your time series.

Illustration: The weather service publishes daily temperature histories for hundreds of locations. Pick a location, such as your home town, and a time period, such as 1905-2005, and fit an ARIMA process.

For a more imaginative student project, compare two adjoining cities with different climates, such as San Francisco and Oakland, or compare a weather station before and after it became a large city.

We have posted Excel data bases of daily temperatures by location. Because we provide so much guidance for this student project, include illustrative worksheets with cell formulas and VBA macros, we expect a good student project on this topic.

Illustration: In past years, we accepted student projects looking at monthly temperature, since it was not easy to get daily temperature for long periods. Now you have all the daily temperature time series you need; don't use monthly temperature.

Take heed: You can do a student project on many other weather topics, such as hourly temperature, rainfall, smog in Los Angeles (or China), by getting other information from weather web sites.

Sales: You can use sales for any industry, including insurance. The web has thousands of sites with business data. A monthly time series shows seasonality, trends, and often business cycles.

Take heed: We show several time series with industry sales (autos, homes, electronics, furniture, department stores, and so forth). You can use any of these time series, a time series from other web sites, or a time series from a private firm.

Birth rates, baby names, marriage rates, divorce rates, mortality rates, and other life events are good time series for the student projects. These topics are fascinating, and the data on the internet are extensive. We show some sample files, such as Excel workbooks from the FBI crime web site.

Take heed: We encourage you to post ideas, links to web sites, and Excel attachments with possible data. These ideas may stimulate student projects by you or other candidates.

Note: Some candidates send us an excess work-book and ask if they can use the data for a student project. We do not answer these questions:

● The answer is almost always yes, if you properly use statistical techniques to analyze the data. We don't want to say "yes" and then receive a student project that does not do the needed work.
● If the answer is "no" because the data have too few observations, the problem is obvious. Don't try to fit 10 years of profits to an ARIMA process.
● If the data contain missing observations or other problems, your write-up explains the problems and your solutions.

You can also use data from your actuarial work, such as personal auto average claim severities. You have these data in Excel work-books, you understand the attributes of the data, and you may already have done much of the work. The student project may be easy to complete. It is fine to use data that you have already collected for other purposes.

Imagination and web surfing give dozens of ideas. Candidates have done student projects on sunspots, airplane crashes, gas prices, border crossing, and other topics.

Illustration: The elections in 2008 will generate hundreds of ideas for student projects, both regression analysis and time series. You can examine the effects of explanatory variables on voting behavior (age, sex, wealthy, religion, ethnic group, education, and so forth).

We post past student projects for many reasons: to illustrate how to apply the statistical techniques, how to adjust for trend, seasonality, cycles, and changes between time periods, and to point out common errors that make your analysis less efficient. The past student projects stimulate ideas for other candidates. Even if the analysis in the student project is not that good, we may post the first paragraph to give ideas to other candidates.

~ Some candidates worry that if they pick their own time series, the ARIMA modeling might not work well.
~ They worry that if they do not fit an ARIMA process that passes Bartlett's test and the Box-Pierce Q statistic, their student project does not pass.
~ They worry that if they choose their own data, they might overlook important effects in their models.
~ They worry that if they pick their own data and no ARIMA process fits well, NEAS will ask them to do the student project again.

The student projects are not intended to produce the true models. Actual statistical work requires extensive analysis of exogenous factors that influence the time series. The statistician spends several days reviewing the relations of the time series and other data.

Do not hesitate to choose other time series or regression data sets. We do not reject student projects because we don't like the choice of topic. If you use statistical techniques properly, your project is fine.

Illustration: A student project on sunspots will not determine the real sunspot process. But sunspots are a good choice for a time series student project. They are a stationary time series or a homogeneous non-stationary time series with several attributes that can be fit to an ARIMA process. The student project selects the time intervals (day, month, or year), analyzes the stationarity or the order of homogeneity, and fits an ARIMA process.

For short time intervals, such as one week, sunspots are clearly autoregressive, since the same sunspot continues for longer than a week. Even over longer periods, sunspots have periods of high or low frequency. The type of ARIMA process is unclear, and you examine graphs, correlograms, trends, cycles, differences, and the autocorrelations.

Some astronomers believe that sunspots have long-term cycles. It is hard to model multi-year cycles with simple ARIMA processes, and we do not expect you to construct an AR(1) or AR(2) model which fits exactly. Your student project would focus on graphing the time series and the correlograms to identify the cycle.

Some candidates want confirmation from NEAS that a particular time series can be used for the student project. Use the following guides:

If the time series is a white noise process,.choose something else. Few time series are pure white noise. You can recognize a white noise process intuitively: random draws from a distribution. If the time series seems like a white noise process, use Bartlett's test or the Box-Pierce Q statistic to test for white noise.

A random walk has first differences that are a white noise process, so random walks are not good for the student project. We don't recommend stock prices for the student project.

Take heed: We don't rule out random walks. You may compare the time series of stock prices from two firms in the same industry, such as two insurers. Your time series might compare the correlation of the two time series.

If the time series has few points or discrete values, choose something else. The time series values do not have normal distributions, and hypothesis testing can not be used (in the form taught in this course). Don't choose earthquake frequency by year in California. This is a white noise process with rare and discrete values.

If the time series has many points and is not a white noise process, it is fine.