Primary research questions:

- (1)How much of IPC data can be predicted with satellite and market data (replicate Erin's paper for Kenya)
- (2) How well does the IPC data map up with household level measures of food insecurity (replicate part 2 of Erin's paper for Kenya)

Secondary goal:

- (3) Create dataset so that county-level yield production functions can be created

Tertiary goal:

- (4) Complement the dataset with other county-level statistics. But as the relationship between climate and welfare outcome variables cannot be causally estimated using these data; I would give this lower priority.

Table: Proposed steps

Timeline	Task	Output	Details
First week of February	Define the empirical specification of the regression	Stata do file	See the paper of Erin.
Second and Third week of February	Data entry of FEWSNET/IPC data	County/monthly data	Download data from FEWSNET and NDMA website and enter manually
Last week of February	Data entry markets	Excel file monthly/market for maize and beans	First check: http://ratin.net/site/page/kenya And the FAO/GIEWS sites Then, do manual entry using the Leading Economic Indicators* (online from 2007 onwards, in library before this date)
Last week of February	Data entry other control variables	Excel file yearly/county	Use online county statistical abstract to select variables and create framework. Do manual data entry using the County Development Profiles* Alternative for the yield variables: -Use NDVI data instead? - Phone ministry of ag once more - Check with Marshall Burke
March	Compile climate variables	GIS daily files?	CHIRPS; check with Jonathan and the global studies team for the best dataset here
April	Merge data		

May-June	Do steps (1) and (2) of the data analysis	

*Manual entry will require a personal visit to the Kenya National Bureau of Statistics. This can be done by Monica (I would propose the last week of Feb as she then will overlap with Jonathan most likely). Or we can hire a local to go. If the latter, we should prepare the exel file with variables ahead of time.

Data available at the Kenya National Bureau of Statistics Library

Goal: Find the right source for county-level data – ideally all in one spot

Method: Very old-fashioned going through every single shelf and picking out each book that looked like it could contain these statistics

- <u>Best source: County Statistical Abstracts</u>. This year is the first year these are being published. They have the following relevant modules: Soils, Rainfall, Retail/Wholesale prices of agricultural inputs, Population, Egg production, Fish caught, Crop Production and Area Cropped, Milk produced, Number of animals and Animals slaughtered, Road network (numbers), Telephone and Postal service, School education (enrolment), Morbidity and death rates, Market prices of food. They have data for 2013, and sometimes also 2014 and previous years. These are also available online at https://www.knbs.or.ke/publications/

  Missing variables: Average levels of education for adults, illiteracy rates for adults (but these could be constructed from survey data if need be), infant mortality more specifically. Also note that not all county books are comparable – while they try to cover the same stats, some cover more than others.

- Socio-economic atlas of Kenya – based on Census data of 2009: County-level statistics on birth/deaths and population. Talking about Census data, there is also the actual census report for each census. This reports all data by district level (which is a lower level than county). The books are also all available in the library. This is also available online.

- Statistical abstract: This is produced every year. The various years are available at the library at the Kenyan Bureau of Statistics; this is available online as well. They have some county level statistics as well, in particular: Population, Birth/Death, Primary and Secondary School Enrolment, Soils.

- Leading economic indicators: This is the coolest thing ever. This is a monthly publication by the Central Bureau of Statistics. I am unsure of the start date – this was an entire bookcase and they were not in order – but a random selection revealed that this seems to go back to the 90s at least. While they focus on macro-economic indicators, they have the market data you are looking for: Monthly market data for the various agricultural outputs (I spotted maize, beans

for 43 market towns). This is available online from 2007 onwards. Copying them for earlier years would likely take about half a day if we were to do this ourselves.

- County development profile. These seem to be the precursor of the County Statistical Abstracts. This is verbal in nature, and harder to extract info from, but seems to have a similar set of variables, and often 4 data points (even though some are projections). This was published in 2013. Unsure whether these are online. The last few pages have some statistical tables that can be used (but probably overlap with the county statistical abstract – I copied these and will take them back to Sussex). Before these county development plans, there were the district development plans. These go back to the 80s and before (as far as I can tell). These are also narrative in nature, but appear also to have all the stats we need. These are not online.

Some logistics to plan a visit there:

- The library is located in the Herufi building of the National Statistics Center. You enter through the armed gates of the treasury where you will need to show a work-ID
- The library is open in the morning and afternoon – a very long lunch break is taken
- While it is possible to photocopy, an elderly gentleman does this for you, at a very slow pace. To photocopy large amounts will take days at his pace; perhaps persuading them to allow you to do the copying yourself? It would be 4 KS per page.
- A better option would be on the spot data entry in excel. There are many spots to sit with a computer. There are no outlets though, so one needs to come with a fully charged computer.
- The WiFI does not work at the library.