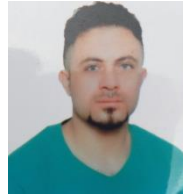# Statistical Analysis of Agricultural Products of Iraq

Alaa Azeez Abed
M.Sc Statistics, Student,
Nationality of Iraq
Department of Statistics
University College of Sciences
Osmania University, Hyderabad, India

ABSTRACT:

Sunflower, Paddy and Dates is the fundamental yield in the pivot framework in Iraq country which is placed in the European piece of Iraq. As this zone is the primary sunflower growing district in Iraq, an extensive review was sorted out there so as to focus the employments of farming assets and their adequacy in sunflower generation in Iraq. Information aggregated through a poll flowed to 182 farming undertakings in Iraq areas in the years 2007-2012 generation year were utilized as a part of the study. The inspected areas were arranged in 3 gatherings as per information utilized for sunflower production of every unit territory. It was discovered that homestead size/land territory was the fundamental creation calculate every 3 gatherings.

The ANOVA test is a test of the null hypothesis that the means of all of the groups are equal, versus the alternative hypothesis that the means are not all equal. The type of ANOVA test we are doing is the simplest ANOVA test, and it is just a generalization of the t-test of difference in means of two groups to the situation in which we have three or more groups. You can think of the two-group situation, in which we apply the t-test for the difference in two means, as a specific case of the more general ANOVA test.

By analyzing the values of sunflower, paddy and dates production in according to area is done by using above mentioned two methods (T-test, ANOVA) and the results obtained will be presented.

1. INTRODUCTION

The Republic of Iraq is in southwest Asia between scopes 29° 5' and 37° 22' N and longitudes 38° 45' and 48° 45' E; it structures the eastern boondocks of the Arab nations. Turkey to the north, Iran to the east, Jordan, Syria and the Kingdom of Saudi to the west, and the Arabian Gulf, Kuwait and the Kingdom of Saudi Arabia to the south, fringe it. Its extraordinary natural, organic and social gimmicks, which are not at all like anyplace else in the Arabian Peninsula, describe the nation. It has an aggregate territory of 438 317 km.

Around 97 percent of the nation is in dry grounds with low and unpredictable precipitation. In many parts horticulture experiences high rates of evapotranspiration that surpass precipitation (Koucher, A. (1999)). Temperature differs generally (10 - 40 °C) amid the developing season, especially in desert areas.

Iraq is an agrarian nation. Albeit a significant bit of its rural terrains are under watering system, regardless it depends exceptionally on rainfed horticulture for grain and sheep creation. Rainfed agribusiness is honed in the northern parts where the mountains, foothills, and

Jazeera desert are placed. These three areas depend for the most part on precipitation for horticultural generation and supply a significant piece of the grain (wheat and grain) devoured in the nation.

Of the 120 000 km2 of cultivable area, which embody 26.4 percent of the aggregate zone, there are 40 000 km2 in the rainfed northern locale and the rest is in the irrigable regions of the Mesopotamian fields.

Rangelands cover 75 percent of Iraq and contribute a substantial offer to the national abundance of the nation and give a work to a large number of individuals through brushing domesticated animals.

These grounds which are inadmissible for cultivating in light of dry atmosphere, poor and rough soils, unpleasant geography and other unfavorable components, make up most of the area in Iraq. The genuine commitment of the rangelands to the sustaining of domesticated animals has not been evaluated, however it has been assessed that range search makes up 90 percent of domesticated animals encourage (Omer, T. M. A. (2011)). It is evaluated that around 7, 900 000 sheep and goats are continued the reach. Rangelands in Iraq have a few insufficiencies for supporting brushing creatures.

Restricted and exceedingly variable precipitation causes wide changes in scrounge creation as far as both amount and quality. Intermittent dry seasons regularly lessen plant development to zero bringing on overwhelming misfortunes and the aftereffect of changing search supplies is reflected in the creatures' wellbeing and gainfulness by moderate development, deferred development, and diminished proliferation. Then again, in the northern and southern Badia (steppe) range gainfulness shifts from 60 to 80 kgDM/ha/year. Amid the brushing season search delivered in these pastures infrequently surpasses the necessities of stationary animals by three times. There are conceivable outcomes to significantly enhance vegetation assets through enhanced administration methods (Omer, T. M. A. (2011)).

Overgrazing is a noteworthy issue. There are no composed endeavors to realize enhanced reach administration to expand domesticated animals creation while keeping up extent assets.

To date data on the nature, degree, present utilization, and creation possibilities of the rangelands is deficient. It is clear that there are numerous and complex issues to be determined before the proceeding with debasement of the rangelands can be stopped and enhancements can be real (Omer, T. M. A. (2011)).

The Iraq country's economy is heavily dependent on agriculture, which accounts for 27% of the GDP and 35% of foreign currency.

The country has a dual agricultural economy (the smallholder sub-sector and the commercial/large-scale sub-sector). The small holder farmers dominating the agricultural sector are estimated to be 31 million. These agricultural holders carry out rain fed agriculture, producing a variety of crops mainly for subsistence purposes. These holders account for most of the food produced in the country.

The main food crops grown in the country are maize, sorghum, millet, cassava, sweet potatoes, bananas, pulses, paddy and wheat. Cash crops grown in Tanzania include coffee, Cashew nut, Tea, cotton, tobacco and sisal. On average the crop sub sector contributes about 34.8% percent of the Agricultural GDP (Mayer, J., Chambers, B., & Farooq, A. (Eds.). (1999)).

The agricultural sector is the main source of employment and livelihood for about 77.5% of the population. It is an important economic sector in terms of food production, employment generation, production of raw material for industries, and generation of foreign exchange earnings (Mayer, J., Chambers, B., & Farooq, A. (Eds). (1999)).

1.2 CLIMATE:
Except for a small part in the northeast, the whole country has an arid or desert, subtropical continental

climate. Summers are hot and completely dry; there are rarely any clouds during four months of summer. Winters are cool. The mean annual rainfall is less than 100 mm in the south-western half of the country increasing to about 140 mm in the centre near Baghdad and to about 200 mm along a line passing about 25 kilometres east of Sinjar in the north, about 40 kilometres west of Mandali and about 80 kilometres east of Baghdad. The 400 mm isohyet passes through Mosul, Erbil, Khanaqin and Mandali. From there to the northeast the rainfall increases gradually to about 1 000 mm in the Zagros Mountains. However, there is great variation in rainfall from year to year especially in the region getting less than 400 mm of rain per year (Armelagos, G. J., & Cohen, M. N. (Eds.). (1984)).

The temperatures are pretty much comparable all through the nation aside from in the mountain territories in the upper east. The month to month mean greatest temperatures for July run from 38 oC at Rutba to 43 oC in Baghdad. The most elevated maxima in June, July, and August range between 43 oC and 50 o 0C. The month to month mean minima for January extend between 1 oC in the south western desert and the north eastern foothills to 8 oC in the focal piece of the waterway plain. The most minimal least is about -14.5 oC in the northern desert, -11 oC in the foothills and -8 oC in the focal piece of the waterway plain. Indeed at Basra close to the coast the least is -4.5 oC, demonstrating the impact of chilly waters (Buringh, P. (1960).

The whole area has a very limited use for grazing of camels, goats, and sheep for two or three months in winter and spring because it is only after the rain that some grasses and shrubs come up and provide grazing for short periods. In some wadis in the north wheat and barley are grown under dry farming systems in years of above average rainfall. Little improvement is possible in the present systems of grazing, or dry farming of wheat and barley in wadis in some parts. It will be only after introducing some organized management system of grazing in the hilly areas of higher rainfall that some improvement of the grazing land may be possible in this area through delaying the grazing to allow the vegetation to develop properly.

1.3 CLIMATE FOR DATES:

The climatic factors influencing date growth and production are discussed in this research. Temperature, rain, humidity, light and wind are the most important climatic factors which determine the suitability of a specific site for growing date palm.

To compare the climates of different places with regard to their suitability for date cultivation, it is appropriate to use the average of the two extremes. However, in order to differentiate between places which may grow dates and those that are not suitable, it is recommended that use be made of the annual average of the daily averages.

Date palm is cultivated in arid and semi-arid regions which are characterised by long and hot summers, no (or at most low) rainfall, and very low relative humidity level during the ripening period. Exceptional high temperatures ($\pm 56°C$) are well endured by a date palm for several days under irrigation. During winters, temperatures below $0°C$ are also endured. The zero vegetation point of a date palm is $7°C$, above this level growth is active and reaches its optimum at about $32°C$; the growth will continue at a stable rate until the temperature reaches $38°C/40°C$ when it will start decreasing.

Even though date palm is a thermopile species, it withstands large temperature fluctuations. Below $7°C$ growth stops and this stage is called a resting period. When the temperature decreases for a certain period to below $0°C$, it causes metabolic disorders which lead to partial or total damage of leaves. At $-6°C$ pinnae margins turn yellow and dry out. Inflorescences are also heavily damaged by frost (Figure 30). When frost periods are suspected, inflorescences should be protected with craft paper bags immediately after pollination (Figure 31). From -9 to $-15°C$, leaves of medium and outside canopy will be damaged and dry out. If these low temperatures are maintained for a

long period (12 hours to 5 days) all leaves will show frost damage and the palm will look as if it was burnt

Date palm culture has mostly been developed in areas with winter rainfall which does not cause harm to the date fruits, but benefits the soils of the plantations by lessivating the deposited surface salt and avoiding the upward movement of salt from lower layers. Table 26 shows that the main date-growing regions are almost rainless until November. In these countries harvest begins mid August until the end of October. Rain during the flowering and harvest season is likely to cause some damage to the fruits.

There is controversy about the effect of rain on pollination and fruit set. Rain that occurs just after pollination is considered as a washing agent that takes away most of the applied pollen.

Another negative effect of rain on fruit set also results from low temperatures that accompany or follow rain. A third factor is the reduction of the flower's receptivity when in contact with water.

Date growers must assume that the rain will affect pollination/fruit set, and any pollination operation immediately preceded or followed by rain (4 to 6 hours) must be repeated.

Rain is also responsible for increasing the relative air humidity creating favourable conditions for cryptogamic diseases that result in the rotting of inflorescences. A light spring shower with a high relative humidity followed by warm temperature just before pollination causes the Khamedj disease (Mauginiella schaettae). This high relative humidity is also associated with the pollen's explosion (Further details on this subject are to be found in the Chapter "Pollination and Bunch Management").

The extension of date palm culture to the southern part of the Sahara Desert (Sahel) came as an exception to the above with heavy rain in July, August and September, and a dry period during October, November, and December.

1.4 NEED OF THE STUDY:
1. To study the production of sunflower and paddy in Iraq.
2. To study the growth of sunflower and paddy crop in Iraq from 2007-2012.

3. To study the importance of increasing the production of sunflower and paddy in Iraq.

1.5 PROBLEM STATEMENT:
The research problem statement is as follows:
1.   To study the growth of sunflower and paddy crop in Iraq from 2007-2012.

2.   The data is collected from different sources.

3.   To analyze and present the production of sunflower and paddy in Iraq.

1.7 PROCEDURES:
The research procedure is as follows:
•   The first stage of the research is about collecting the data from different sources.

•   The second stage to study is about the study on production of sunflower and paddy in Iraq from 2007-2012.

•   The third stage of the research is statistical analysis and presenting the same.

•   The final stage is presenting the conclusions of the research.

1.8 LIMITATIONS:
The research has some limitation as follows:
•   The research limits in analyzing the exact production rate of sunflower and paddy in Iraq.

•   The data collection for the research may limits in acquiring bulk data.

2. LITERATURE REVIEW

As per Sacks, W. J., Deryng, D., Foley, J. A., & Ramankutty, N. (2010), exploration investigate spatial connections between planting date and atmosphere for two of them, and contrast our investigation and an audit of the writing on components that drive choices on planting dates.

We digitized and geo-referenced existing information on product planting and collecting dates from six sources. We then inspected connections between planting dates and temperature, precipitation and potential evapotranspiration utilizing 30-year normal climatologies from the Climatic Research Unit, University of East Anglia (CRU CL 2.0).

We show worldwide planting date designs for maize, spring wheat and winter wheat (our full, openly accessible information set contains planting and reaping dates for 19 noteworthy harvests). Maize planting in the northern mid-scopes for the most part happens in April and May. Every day normal air temperatures are normally c. 12–17 °C at the time of maize planting in these areas, despite the fact that dirt dampness regularly decides planting date more straightforwardly than does temperature. Maize planting dates differ all the more generally in tropical areas. Spring wheat is normally planted at cooler temperatures than maize, between c. 8 and 14 °C in calm districts. Winter wheat is by and large planted in September and October in the northern mid-scopes.

In calm locales, spatial examples of maize and spring wheat planting dates can be anticipated sensibly well by accepting an altered temperature at planting. Be that as it may, planting dates in lower scopes and planting dates of winter wheat are more hard to anticipate from atmosphere alone. Partially this is on account of planting dates may be decided to guarantee a positive atmosphere amid a basic development stage, for example, blossoming, as opposed to guarantee an ideal atmosphere ahead of schedule in the crop's development. The absence of consistency is additionally because of the pervasive impact of innovative and financial variables on planting dates.

Agreeing Monfreda, C., Ramankutty, N., & Foley, J. A. (2008), examination Croplands spread ~15 million km2 of the planet and give the majority of the sustenance and fiber fundamental to human prosperity. Most worldwide area spread information sets from satellites bunch croplands into simply a couple of classes, in this way barring data that is discriminating for noting key inquiries running from biodiversity protection to nourishment security to biogeochemical cycling. Data about horticultural area utilization practices like product choice, yield, and compost utilization is significantly more restricted. Here we present area use information sets made by consolidating national, state, and region level statistics insights with an as of late overhauled worldwide information set of croplands on a 5 min by 5 min (~10 km by 10 km) scope longitude network. The subsequent area use information sets portray around the year 2000 the region (reaped) and yield of 175 particular yields of the world. We total these individual harvest maps to create novel maps of 11 noteworthy yield gatherings, product net essential generation, and four physiologically based harvest sorts: annuals/perennials, herbaceous/bushes/trees, C3/C4, and leguminous/nonleguminous.

Ainsworth, E. A. (2008), research says Rice is arguably the most important food source on the planet and is consumed by over half of the world's population. Considerable increases in yield are required over this century to continue feeding the world's growing population. This meta-analysis synthesizes the research to date on rice responses to two elements of global change, rising atmospheric carbon dioxide concentration ($[CO2]$) and rising tropospheric ozone concentration ($[O3]$). On an average, elevated $[CO2]$ (627 ppm) increased rice yields by 23%. Modest increases in grain mass and larger increases in panicle and grain number contributed to this response. The response of rice to elevated $[CO2]$ varied with fumigation technique. The

more closely the fumigation conditions mimicked field conditions, the smaller was the stimulation of yield by elevated [CO2]. Free air concentration enrichment (FACE) experiments showed only a 12% increase in rice yield.

The rise in atmospheric [CO2] will be accompanied by increases in tropospheric O3 and temperature. When compared with rice grown in charcoal-filtered air, rice exposed to 62 ppb O3 showed a 14% decrease in yield.

Many determinants of yield, including photosynthesis, biomass, leaf area index, grain number and grain mass, were reduced by elevated [O3].

While there have been excessively few investigations of the connection of CO2 and O3 for meta-examination, the communication of temperature and CO2 has been mulled over all the more generally. Lifted temperature medications invalidated any improvement in rice yield at raised [CO2], which proposes that distinguishing high temperature tolerant germplasm will be vital to acknowledging yield advantages later on.

As per Biglari, F., AlKarkhi, A. F., & Easa, A. M. (2008), Edible parts of date palm (Phoenix dactylifera) natural products (DPF) from Iran were investigated for their cell reinforcement exercises (AA) utilizing Trolox proportionate cancer prevention agent limit (TEAC) system, 2,2′-azinobis (3-ethylbenzothiazoline-6-sulphonic corrosive) radical cation (ABTS +) examines and the ferric diminishing/cell reinforcement power strategy (FRAP test). The aggregate phenolic substance (TPC) and aggregate flavonoid content (TFC) of the DPF were measured utilizing Folin–Ciocalteau and aluminum chloride colorimetric techniques, individually.

The specimens utilized incorporated four sorts of delicate dates (SD) specifically Honey date, Bam date, Jiroft date and Kabkab date; three sorts of semi-dry dates (SDD) in particular Sahroon date, Piarom date and Zahedi date and one kind of dry date (DD) which was Kharak date. The AA (ABTS test) of the DPF

were 22.83–41.17, 47.6–54.61 and 500.33 µmol Trolox equivalents/100 g dry weights (dw) for SD, SDD and DD, separately. The AA (FRAP test) every 100 g dw test were 11.65–20, 19.12–29.34 and 387.34 µmol FRAP for SD, SDD and DD, separately. The TPC ran from 2.89 to 4.82, 4.37 to 6.64 and 141.35 mg gallic corrosive equivalents (GAE)/100 g dw, while TFC extended from 1.62 to 3.07, 1.65 to 4.71 and 81.79 mg catechin equivalents (CEQ)/100 g dw test for SD, SDD and DD, individually.

Relationship investigations showed that there was a straight relationship in the middle of AA and the TPC or TFC of DPF. This work exhibits the capability of Iranian dates as cancer prevention agent food ingredients.

3. METHODOLOGIES
3.1 MATERIAL/INSTRUMENTS:
The data collected as Agri products of Iraq has analysed by Analysis of Variance and the students t-test to know that whether the yield in various locations of crops during the 2007-2012 is homogeneous for the Agricultural Products of Iraq and t-test was employed to know pair wise differences occur or not is being tested for the yield.

3.2 T-TEST AND ANOVA
- ANOVA is better understood as analysis of means. The methods employ ratios of variances in order to establish whether the means of several groups differ.
- An independent sample t test, in which only two group means are compared, is a special case of one-way ANOVA.
- One-way ANOVA simply means that there is only one explanatory variable, while two-way ANOVA implies two explanatory variables.
- The problem of ANOVA is to determine whether there is a statistically significant difference between the means, relative to the spread of the observations in each group.

- If you remember, variance (i.e. the square of standard deviation) is a measure of spread. Hence, this is why the method called ANOVA.

3.2.1 One-Sample t test

- It is a test to evaluate whether the mean of a test variable is significantly different from a test value.

- For instance, we draw a sample from your class and test whether your class quiz score mean is different from 50, presuming that 100 is the full score.

- Thus, the hypotheses involved are:

  H_0: $\mu - 50 = 0$

  H_A: $\mu - 50 > 0$ or < 0 (one-tailed)

  H_A: $\mu - 50 \neq 0$ (two-tailed)

- In here, only one variable is being tested. Even when the mean score is statistically significant, i.e. significantly different from 50, how can we assess the effect, a measure similar to the measures of association?

- We can use d to evaluate the degree that the mean scores on the test variable differ from the test value in standard deviation unit.

- The formula is:

$$d = \frac{t}{\sqrt{N}}$$

Where t is the t value and N is the sample size.

- t is calculated from the formula:

$$t = \frac{\bar{x} - a}{s / \sqrt{N}}$$

Where $\bar{x}$ is the sample mean, s is the sample standard deviation, and a is the tested value with sample size N

- by substituting t into d, it is found that d also equals to "mean difference divided by standard deviation" or $d = \dfrac{\bar{x} - a}{s}$

- d values of .2, .5 and .8, regardless of sign, are by convention interpreted as small, medium, and large effect sizes respectively.

- Note that the shape of the t distribution depends on degree of freedom. df = N - 1

3.2.2 Paired-Samples t test

- Each case must have scores on two variables for a paired-sample t test. The paired-sample t test evaluates whether the mean of the difference between these two variables is significantly different from zero.

- Usually, it is used for a repeated-measures design in which a participant is assessed on two occasions or under two different conditions on one measure.

- Or in a match-subjects design, participants are paired and each participant in a pair is assessed once on a measure.

- A commonly acceptable sample size to yield accurate p values is 30 pairs of scores. Larger sample sizes may be required to produce relatively valid p values if the population distribution is substantially non-normal.

- An example: after the first quiz in Social Research, I deliberately scare you about the "poor" results so as to motivate you to do the second quiz better. I want to know whether there is a scaring effect by testing the first and second quiz scores for each of you.

- The hypotheses are:

  H_0: $\mu_{difference} = 0$

  H_A: $\mu_{difference} > 0$ or < 0 (one-tailed)

  H_A: $\mu_{difference} \neq 0$ (two-tailed)

- To calculate sample mean difference ($\bar{x}_{difference}$), one needs to calculate the difference between a pair of scores, sum the differences up, and then take the average.

- t is calculated as follow:

$$t = \frac{\bar{x}_{difference} - 0}{s_{difference} / \sqrt{N}}$$ where $\bar{x}_{difference}$ is the sample mean of score differences, $s_{difference}$ is the sample standard deviation of score differences, and N is the sample size.

- Note that the shape of the t distribution depends on degree of freedom.

  df = number of pairs – 1

- Again, d is used to assess the effect, which has formula $d = t/\sqrt{N}$

3.2.3 Independent Samples t test

- The independent-samples t test evaluates the difference between the means of two independent groups.

- The statistical null hypothesis states that the two population means μ_1 and μ_2 are equal. Or put it differently:

  H_0: μ_1 - μ_2 = 0

  H_A: μ_1 - μ_2 > 0 or < 0 (one-tailed)

  H_A: μ_1 - $\mu_2 \neq 0$ (two-tailed)

- Two assumptions of the t test are:

1. the populations sampled are normally distributed

2. the population variances are equal (even when the variances are not equal, SPSS will also give an independent-samples t test without the assumption of equal variances)

- The test statistic t is calculated as follow:

- The numerator of the t is simply the mean difference between the two groups.

- The denominator of it is quite complicated. It should be the common population variance. However, since this parameter is unknown to us.

- Hence, we use the weighted average of the two sample variances of the two groups to estimate the population variance.

- The weights we give to the two groups are simply their respective number of sample cases less one divided by the total number of cases less two. It is because we believe that the more cases a sample has, the larger variance of the sample will be.

- The t-value is computed by the following equation:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

where $s^2 = \dfrac{n_1 - 1}{n_1 + n_2 - 2}s_1^2 + \dfrac{n_2 - 1}{n_1 + n_2 - 2}s_2^2$

- $n_1 + n_2 - 2$ is the degree of freedom of the t-value. The degrees of freedom for a particular sum of squares is equal to the number of terms in the sum we need to know in order to find the remaining terms and thereby complete the sum.

- In SPSS the test of equal variances is known as Levene's test. If the Levene's test is significant, that means the variances are not equal. Then you should look at the row with "equal variance not assumed". Otherwise, look at the other row.

- Again, d is used to assess the effect, which has formula $d = t\sqrt{\dfrac{N_1 + N_2}{N_1 N_2}}$ where N_1 and N_2 are the sample sizes of the two groups.

3.2.4 SPSS Examples

- All t tests can be found by clicking Analyze and then Compare Means.

  Let take an example of independent-samples t test: we want to know whether gender has an effect on one's educational achievement. In the 1991 US general social survey dataset, gender is labeled as sex and education as educ measuring the highest year of school completed.

- Here are the steps:

1. Click Analyze, click Compare Means, then click Independent-Samples T Test

2. Choose educ, and click ▶ to move it to the Test Variable(s) box.

3. Choose sex, and click ▶ to move it to the Grouping Variable box.

4. Click the Define Groups option and then enter 1 and 2 into Group 1 and Group 2 boxes. (in variable sex, 1 is the value label for males and 2 is for females)
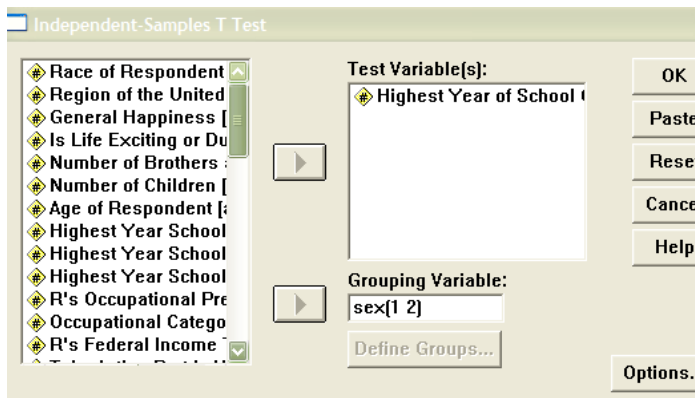
5. Click Continue and OK.

3.2.5 Output interpretation:

- From the Group Statistics table, we know that males have more education than females. But whether is the difference statistically significant, not produced by chance?

Group Statistics

	Respondent's Sex	N	Mean	Std. Deviation	Std. Error Mean
Highest Year of School Completed	Male	633	13.26	3.149	.125
	Female	877	12.63	2.829	.096

- To answer the question, we have to look at the Independent Samples Test table. First we need to determine whether the equal variances condition is assumed by looking at the Levene's test.
- If the Levene's test is significant at .05 level, we should look at the row with equal variance not assumed. Otherwise look at the equal variance assumed row.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Highest Year of School Completed	Equal variances assumed	11.226	.001	3.887	1508	.000	.60	.155	.298	.906
	Equal variances not assumed			3.824	1276.454	.000	.60	.157	.293	.911

In the Levene's test, the sig. is .001 which is < .05. Thus we can rejected the equal variances null hypothesis. We should choose the row "equal variances not assumed" and look for t, which is 3.824 and significant at .001 level.

- Since the p-value of the t statistic is .000, it is sure less than .05, .01 and .001 significance levels. Thus we can conclude that gender does have an effect on one's educational achievement with male having the advantage.

- By using the formula of $d = t\sqrt{\dfrac{N_1 + N_2}{N_1 N_2}}$ or

 $3.824\sqrt{\dfrac{633 + 877}{633 * 877}}$, we get the value of d as 0.20 which indicates only a weak relationship between gender and education.

- Let's take another example, about paired-samples t-test.
- If we want to see whether one's parents' educations are matched, we can use the paired-samples t test. In other words, we want to test the differences between the father's and mother's education of respondents.
- Here are the steps:
1. Click Analyze, click Compare Means, then click Paired-Samples T Test
2. Choose paeduc and maeduc together, and click ▶ to move it to the Paired Variables box.
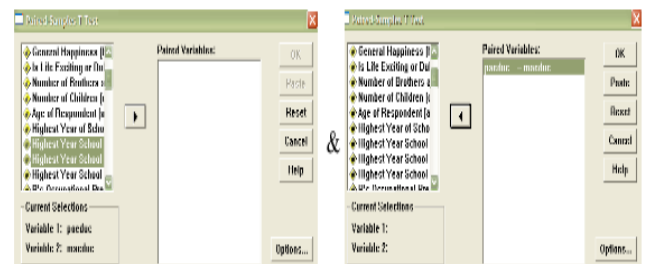
3. Click Continue and OK.

3.2.6 Output interpretation:

- From the table Paired Samples Test, the Sig is .943 which is too large to be significant. Moreover, from the 95% confidence interval of the difference, which ranges from -.20 to .19, zero is included in the interval. That means we cannot reject the null hypothesis that there is no difference between father's and mother's education.

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Highest Year School Completed Father - Highest Year School Completed Mother	-.01	3.115	.100	-.20	.19	-.072	975	.943

- Since the result is not significant, we don't need to calculate d.

3.3 ONE-WAY ANOVA

- In one-way ANOVA, each case must have scores on two variables: a factor (or independent variable) and a dependent variable. The factor divides cases into two or more groups or levels, while the dependent variable differentiates cases on some quantitative dimension. The ANOVA F test assesses whether the group means on the dependent variable differ significantly from each other.

- If the factor divides cases into two groups, the analysis is essential similar to an independent-samples t test. Thus, you can say that one-way ANOVA is to compare the means of more than two groups.

- One-way ANOVA is based upon the assumption that we can decompose each observation into three additive terms:

- observation = overall mean + deviation of group mean from overall mean + deviation of observation from group mean

- The above equation can be interpreted as follow:

- observed value of Y = constant + effect of being in a particular group (effect of the factor X) + effect of all other variables (residual)

- The summation of "the square of the deviation of each group mean from the overall mean times the number of cases in the corresponding group" is known as "between group sum of squares" (BSS).

- The summation of "the square of the deviation of each observation from the corresponding group mean" is known as "within group sum of squares" (WSS). It is also known as residual sum of squares (RSS).

- The degree of freedom of the BSS equals to the number of groups less one (df_{BSS}).

- The degree of freedom of the WSS equals to the number of cases less the number of groups (df_{WSS}).

- The statistical hypothesis of ANOVA is that all the group means are equal, i.e. there is no difference between groups:

- H_0: $\mu_1=\mu_2=\mu_3=\dots=\mu_n$

- H_A: any two group means are unequal

- In order to test the above hypothesis, we use the F-test. The F-ratio is computed as:

$$F = \frac{BSS / df_{BSS}}{WSS / df_{WSS}} = \frac{Mean_{BSS}}{Mean_{wss}}$$

- The BSS tells us how large the effect of the factor or explanatory variable (i.e. the groups) is on the dependent variable, while the WSS indicates the random variation of the dependent variable due to other uncontrolled variables.

- Put it differently, F-ratio = variation between the means/unexplained (error) variation.

- If in the population all the group means are equal, then the observed differences of the sample group means must be due to random sampling error. Hence, the $Mean_{BSS}$ should approximately equal to the $Mean_{WSS}$. In other words, the F-ratio will be roughly one.

- We can check whether the F-ratio is statistically significant by looking at the F-table with the degrees of freedom, df_{BSS} and df_{WSS}.

- The explanatory power of the independent variable is known as Eta squared (η^2), a PRE

measure, which can be computed as:

$$\eta^2 = \frac{TSS - RSS}{TSS} = \frac{BSS}{TSS}$$

- It should be noted that when we do an ANOVA with only two groups, the F-ratio is equivalent to the square of the t-value.

- In one-way ANOVA, normally it is assumed that the group variances are equal. When the variances are not equal across groups, the results of the ANOVA Table are questionable. The Welch and Brown-Forsythe statistics are alternatives to the usual F test in such a case.

- In order to decide whether to use the Welch and Brown-Forsythe statistics, we can first eye-inspect the standard deviation of each group to see whether there is a large range, then we can perform the homogeneity test (using Levene statistic as in the independent-samples t test) to see whether we can reject the null hypothesis of equal variances.

- If the F test shows that the result is significant, then we need to perform a follow-up test (known as post hoc multiple comparisons) to assess which group means are significant different because our alternative hypothesis is not implying all the group means are different, but at least two of them.

- There are at least two types of follow-up tests: whether to assume equal variances.

- If equal variances are assumed, use Tukey or R-E-G-W Q, otherwise use Dunnett's C. Indeed, there are a number of statistics you can choose from the SPSS menu.

- Through these statistics, we exactly know which pairs of group means are different.

3.4 TWO-WAY ANOVA

- In a field experiment we have learnt that the interaction effects in factorial design. In order to examine the interaction effects, we need to do a two-way or higher level ANOVA.

- With a two-way ANOVA, each participant must have scores on three variables: two factors and a dependent variable. Each factor divides cases into two or more levels, while the dependent variable is a continuous variable.

- F tests are performed on the main effects for the two factors and the interaction between the two factors. Post hoc tests may be used to assess specific hypotheses if main effect tests, interaction tests, or both are significant.

- For a three-way ANOVA, tests are conducted on the three main effects for the factors, the three two-way interactions between pairs of factors, and the one three-way interaction among the three factors.

- Since the statistical techniques used for two-way or higher level ANOVA are beyond what I intend to teach you, I will just give you some brief ideas. For details (only for AP students), you may learn it when you study the degree courses.

- A two-way ANOVA tests three hypotheses:
  The means for the levels of the first factor (or independent variable) are different

  The means for the levels of the second factor (or independent variable) are different

  The pattern of means for one factor differs between the levels of the other factor (testing interaction effects)

- Just like one-way ANOVA, the total variance of the dependent variable is partitioned into different components:
  Total Variance = Between Group Variance + Within Group Variance (error or residual)

- In one-way ANOVA, the between group variance is simply due to the factor (treatment or independent variable) effect. But, in two-way ANOVA:
  Between Group Variance = Treatment 1 + Treatment 2 + Interaction between Treatments

- In other words, the between group sum of squares (BSS) can be decomposed into three sources:

$BSS = BSS_1 + BSS_2 + BSS_{1*2}$ (BSS for treatment 1, 2, and their interaction respectively)

- For each type of BSS, we can divide it by its df and calculate their respective Mean BSS, and then perform the F test.
- The total degrees of freedom are one fewer than the number of cases. The df for each treatment are one fewer than the number of levels in that treatment. The df for interaction is calculated by multiplying the df for each of the treatments. The df for residual are the number of df which are left once all the treatment dfs have been removed from the total. (The dfs are used to do F tests. Indeed, SPSS will do all the calculations for you.)
- From F tests, we will know which components are significant.

3.5 STATISTICAL ANALYSIS:

Generally speaking, so-called linear statistical models state that your outcome of interest (or a mathematical transformation of it) can be predicted by a linear combination of explanatory variables, each of which is multiplied by a parameter (sometimes called a coef fi cient and often denoted b). To avoid having the outcome be estimated as zero if all explanatory variables are zero, a constant intercept (often denoted b 0) is included.

The outcome variable of interest is often called the dependent variable, while the explanatory variables that can predict the outcome are called independent variables.

The terminology in statistics and experimental design may sometimes be somewhat confusing. In all practical applications, models like linear regression, analysis of covariance (ANCOVA), analysis of variance (ANOVA), or general linear models (GLM) are very similar. Their different terminology is due as much to the historical tradition in statistical science as to differences in methodology. Many of these models with their different names and terminologies can be expressed within the framework of generalized linear models.

It was common to develop mathematical methods to estimate parameter values and p -values that could be calculated manually by hand and with the help of statistical tables. Most graduates in statistics are familiar with such methods for simple regression and ANOVA methods. However, recent innovations in mathematical statistics, and not least computers and software, have in an applied sense replaced such "manual" methods. These computer-assisted methods are usually based on the theory of so-called likelihood functions and involve finding their maximum values by using iterations. In other words, these are methods where computer software is needed for most applied circumstances. The theory behind maximum-likelihood estimations is covered in several of the more advanced recommended textbooks.

Linear statistical models are often described within the framework of generalized linear models. The type of model is determined by the properties of the outcome variable. A dependent variable with continuous data is usually expressed with an identity link and is often referred to by more traditional terms such as linear regression or analysis of variance.

If the dependent variable is binary, then it is usually expressed by a logit link and is often referred to by the more traditional term logistic regression. Count data use a log link and the statistical model is traditionally referred to as Poisson regression (e.g., Dobsen and Barnett 2008).

3.6 T-TEST

The t-test, also known as the student t test, is a test of significance that can be used to determine whether a significant difference exists or does not exist between two groups.

There are three kinds of t-tests:

1. t-test for one Sample

2. t-test for Dependent Samples

3. t-test for Independent Samples

When a t-test is used to determine whether the two sample means of two independent samples come from the same population, we use the statistical test called the t-test for Independent Means. This is the most common t-test used in science. The formula for calculating "t" depends upon whether the two samples being compare have equal variances. The null and alternative hypotheses for this type of test are:

H_0: $m_1 = m_2$; the means are equal

H_1: $m_1 </> m_2$; the means are different

This is a two tailed test because the Null Hypothesis does not specify a direction, only the condition of equality.

The assumptions are:

1. The data are Normal

2. The two samples come from distributions that may differ in their mean value, but not in the standard deviation

3. The observations are independent of each other.

Calculate a t-test value, and compare the value with a critical value of t (given in a book). If the t value calculated from the data is equal to or larger than the critical value, you reject the Null hypothesis of H_0: $m_1 = m_2$.

Before we determine the calculated t (t_c), look at the variances. If the variances are equal, calculate t using the formula below:

$$t = \frac{|m_1 - m_2|}{\sqrt{\frac{(n_1 - 1)s^2_1 + (n_2 - 1)s^2_2}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

If the variances are unequal, calculate t as follow:

$$t = \frac{|m_1 - m_2|}{\sqrt{\frac{s^2_1}{n_1} + \frac{s^2_2}{n_2}}}$$

m = mean

n = sample size

s^2 = variance

Using the data below:

State your hypotheses; Set a probability level, called the alpha level $\square\square\square$ = .05; Calculate 't' ; compare the numerical value calculated from the statistical test to a critical value or tabled value (t_t) (obtained from a table/book). In this case $t_t = 2.000$

Reject the Null hypotheses if your calculated value is greater than or equal to the critical or tabled value.

3.7 CALCULATING THE MEAN AND STANDARD DEVIATION

The mean and standard deviations are the most widely used statistical measures. They are used to determine the central tendency (typical score) and variability (spread) of interval data. Mathematically, the mean is the arithmetic average of a number of scores. To obtain the mean, add your scores and divide by the number of scores that you have. Consider the following data, which represents the time needed to complete a reading task, as an example.

Time in Minutes   6   3   5   5   2   7
6   4   3   2   Total =   43

The mean is the sum of scores divided by the number of scores, mathematically:

$$\text{Mean} = \Sigma X/N = 43/10 = 4.3$$

The <u>standard deviation</u>, abbreviated as <u>s</u>, is a measure of the variability of interval data.

The formula for the standard deviation is given below.

$$S = \sqrt{\frac{\Sigma x^2 - \frac{(\Sigma x)^2}{N}}{(N-1)}}$$

To find the standard deviation, subtract the mean from each of the scores, square the deviation, and then add up the squared deviations. This process is outlined below.

Time Scores	Mean	Score - mean	(Score - mean)
6	4.3	1.7	2.89
3	4.3	-1.3	1.69
5	4.3	0.7	0.49
5	4.3	0.7	0.49
2	4.3	-2.3	5.29
7	4.3	2.7	7.29
6	4.3	1.7	2.89
4	4.3	-0.3	0.09
3	4.3	-1.3	1.69
2	4.3	-2.3	5.29

$$\text{Total} = 28.10$$

Therefore, the standard deviation becomes:

$$S = \sqrt{\frac{28.10}{9}} = \sqrt{3.12} = 1.77$$

For Grouped Data

Often data will be reported in terms of grouped observation and the calculation of the mean and standard deviations are modified somewhat. Although the logic of the computations and interpretations are not changed, the computational procedures become more tedious. Consider the following example of age classifications:

Ages	f
51 - 60	3
41 - 50	10
31 - 40	15
21 - 30	11
11 - 20	5

To calculate the mean and standard deviation you must determine the midpoint for each of the groups of observations. Adding the upper and lower scores for each interval and dividing by two can obtain the midpoints. For example, for the first group of data the midpoint would be (51 + 60)/2 = 111/2 = 55.5. I have redrawn the data below with the midpoints inserted. Also, I have included in the redrawn data a column headed by the term fxMidpoint, which is simply the midpoint multiplied by the frequency

Ages	f	Midpoints	fxMidpoint
51 - 60	3	55.5	166.5
41 - 50	10	45.5	455.0
31 - 40	15	35.5	532.5
21 - 30	11	25.5	280.5
11 - 20	5	15.5	77.5

Total = 1512.0

The mean becomes the sum of the scores in the fxMidpoint column divided by the sample size. The sample size can be determined by adding the f column (N = 44). Therefore, the mean = 1512.0/44 = 34.36 (rounded to 34.4).

To determine the standard deviation we will need to add one additional column to the table of calculations above. This additional column is fx midpoint and I have redrawn our table below with the added column.

Ages	f	Midpoints	fxMidpoint	fxMidpoint
51 - 60	3	55.5	166.5	9240.75
41 - 50	10	45.5	455.0	20702.50
31 - 40	15	35.5	532.5	18903.75
21 - 30	11	25.5	280.5	7152.75
11 - 20	5	15.5	77.5	1201.25

Totals = 1512.0   57201.00

The standard deviation is obtained by a slightly different formula, which is easier to apply in this situation. (You should remember that the new formula gives the same answer as the formula I presented above. It is not too difficult a task to show that the formulas are mathematically the same.) The new formula is presented below:

$$S = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{N}}{(N-1)}}$$

$$S = \sqrt{\frac{57201.0 - \frac{1512.0^2}{44}}{43}} = \sqrt{\frac{57201 - 51957.82}{43}} = \sqrt{121.93} = 11.04$$

4. RESULTS

In this chapter we present the results of the data collected and analyzed

4.1 SAMPLE:

Gover norate	Grand Total	2012 Dates production by kind and governorates for (production tons)						
		Others	Deary	Hillawi	Khadrawi	Sayer	Khistawi	Zahdi
Kirkuk	115	45	2	1	16	–	3	48
Diala	83792	15165	4986	256	2048	3499	10511	47327
AL-Anbar	43196	1160	14	11	164	31	7568	34248
Baghdad	93782	22790	323	274	1717	248	11467	56963
Babylon	100343	7625	1829	219	1286	593	12710	76081
Kerbela	72217	5546	320	61	518	98	9534	56140
Wasit	43649	10042	370	565	1736	285	6588	24063
Salah Al-Deen	19734	5055	67	504	368	188	2756	10796
AL-Najaf	32304	2590	246	424	1466	791	3724	23063
AL-Qadisiya	31436	3574	293	380	916	689	1357	24227
AL-Muthanna	25275	7589	1739	968	2027	1366	1508	10078
Thi Qar	40086	9595	4541	582	10193	5954	1421	7800
Mysan	8967	1609	612	135	2320	1902	656	1733
Basrah	60554	17617	4487	19803	4366	9008	1684	3589
Total	655450	110002	19829	24183	29141	24652	71487	376156

Figure 4.1: Dates production 2012

Cultivated area comparing to production and average yield of Paddy and Sun Flower at Iraq level for (2007-2012)

Table (1)

Details	Crop		Years
	Sun Flower	Paddy	
Total area (100) Donum	279	4974	2007
	193	3390	2008
	168	2197	2009
	226	1919	2010
	179	2638	2011
	87	3188	2012
Production (100) Ton	125	3928	2007
	96	2482	2008
	52	1731	2009
	75	1558	2010
	71	2351	2011
	42	3613	2012
Average Yield per area (KG/Donum)	449.9	789.8	2007
	497.8	731.9	2008
	310.7	787.6	2009
	331.6	812.1	2010
	397.4	891.2	2011
	484.5	1133.5	2012

Figure 4.2: Cultivated area comparing to production and average yield of Paddy and Sun Flower at Iraq level for (2007-2012)

4.2 NEED OF THE STUDY:

- To study and statistical analyze the sunflower and paddy, dates crop growth in Iraq from 2007-2012.
- To collect the data from different sources.
- To analyze and present the production of sunflower and paddy, dates in Iraq.
- For carrying the test we are using two methods, t-test and ANOVA.
- We use t-test for the comparative analysis of the sunflower and paddy production.
- We use ANOVA for carrying the homogenous comparison for different kinds of dates production

4.3 HYPOTHESIS:

The research is mainly focuses on the statistical analyze the sunflower and paddy, dates crop growth in Iraq from 2007-2012. The study needs to evaluate the Production of sunflower and paddy, dates in Iraq. To study the GDP of Iraq based on sunflower and paddy production. We carry here two methods of analysis 1. T-test method, 2. ANOVA method.

T-test method is used to analyse the sunflower and paddy production and in-turn compare them with there are and production value.

ANOVA method is used to make homogenous comparison for different kinds of dates production.

We have two hypotheses for t-test here:

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$

With confidence level $\alpha = 0.05$

CONCLUSIONS:

We have analyzed and present the production of sunflower and paddy, dates in Iraq. For carrying the test we are using two methods, t-test and ANOVA. We use t-test for the comparative analysis of the sunflower and paddy production. We use ANOVA for carrying the homogenous comparison for different kinds of date's production.

The results obtained after the analysis of t-test and ANOVA are present in Chapter-IV.

The total production of different types of dates in iraq is seen in the above table.

MEAN:

The maximum mean is observed in following kind of dates:

- Others (7857.2857),

- Khadrawi (2081.5000)

- Khistawi (5106.2143)

- Zahdi (26868.2857)

The minimum mean is observed in the following kind of dates:

- Deary (1416.3571

- Hillawi (1727.3571)

- Sayer (1760.8571)

STANDARD DEVIATION:

The maximum standard deviation is observed in the following kind of dates:

- Others (6701.15621)

- Hillawi (5209.23689)

- Khistawi (4480.10611)

- Zahdi (23992.99379)

The minimum standard deviation is observed in the following kind of dates:

- Deary (1855.71237)

- Khadrawi (2587.28161)

- Sayer (2668.22874)

The overall total of the total mean resulted in the ANOVA test is 6688.2653, and the total standard deviation resulted in the ANOVA test is 12855.88326.

As per the table 'tests of between-subjects effects' we conclude that, the p-value of significance for Corrected Model, Intercept, kind of dates is lower than confidence level which we assumed in the hypothesis (.000000 < 0.05), then we consider that these has no significant main effect.

The p-value of significance for Governorate is higher than the confidence level which we assumed in the hypothesis (.129 > 0.05), we consider that Governorate has significant main effect.

References

- Ainsworth, E. A. (2008). Rice production in a changing climate: a meta-analysis of responses to elevated carbon dioxide and elevated ozone concentration. Global Change Biology, 14(7), 1642-1650.

- Alami, Y., Achouak, W., Marol, C., & Heulin, T. (2000). Rhizosphere Soil Aggregation and Plant Growth Promotion of Sunflowers by an Exopolysaccharide-Producing Rhizobiumsp. Strain Isolated from Sunflower Roots. Applied and environmental microbiology, 66(8), 3393-3398.

- Anderson, D., Sweeney, D., Williams, T., Camm, J., & Cochran, J. (2013). Statistics for Business & Economics. Cengage Learning.

- Armelagos, G. J., & Cohen, M. N. (Eds.). (1984). Paleopathology at the Origins of Agriculture (pp. 235-269). Academic Press page (20 – 40).

- Biglari, F., AlKarkhi, A. F., & Easa, A. M. (2008). Antioxidant activity and phenolic content of various date palm (Phoenix dactylifera) fruits from Iran. Food chemistry, 107(4), 1636-1641.

- Brockwell, P. J. (2002). Introduction to time series and forecasting (Vol. 1). Taylor & Francis.

- Brockwell, P. J., & Davis, R. A. (2009). Time series: theory and methods. Springer Science & Business Media.

• Buringh, P. (1960). Soils and soil conditions in Iraq. Baghdad: Ministry of agriculture.

• Carey-Smith, T., Sansom, J., & Thomson, P. (2014). A hidden seasonal switching model for multisite daily rainfall. Water Resources Research, 50(1), 257-272.

• Casdagli, M. (1989). Nonlinear prediction of chaotic time series. Physica D: Nonlinear Phenomena, 35(3), 335-356.

• Chatfield, C. (2013). The analysis of time series: an introduction. CRC press.

• Chatfield, C. (2013). The analysis of time series: an introduction. CRC press.

• Clawson, M., Landsberg, H. H., & Alexander, L. T. (1971). The agricultural potential of the Middle East. The agricultural potential of the Middle East.

• Denison, R. F., Bryant, D. C., & Kearney, T. E. (2004). Crop yields over the first nine years of LTRAS, a long-term comparison of field crop systems in a Mediterranean climate. Field Crops Research, 86(2), 267-277.

• Desikan, P., & Srivastava, J. (2014). Time series analysis and forecasting methods for temporal mining of interlinked documents. Department of Computer Science, University of Minnesota, www-users. cs. umn. edu/~ desikan/publications/TimeSeries. doc, accessed on January, 26.

• Desikan, P., & Srivastava, J. (2014). Time series analysis and forecasting methods for temporal mining of interlinked documents. Department of Computer Science, University of Minnesota, www-users. cs. umn. edu/~ desikan/publications/TimeSeries. doc, accessed on January, 26.

• Franses, P. H. (1998). Time series models for business and economic forecasting. Cambridge university press.

• Furrer, E. M., & Katz, R. W. (2007). Generalized linear modeling approach to stochastic weather generators. Climate Research, 34(2), 129.

• Granger, C. W. (1966). The typical spectral shape of an economic variable. Econometrica: Journal of the Econometric Society, 150-161.

• Hafez, A. (2001, March). Association mining of dependency between time series. In Aerospace/Defense Sensing, Simulation, and Controls (pp. 291-301). International Society for Optics and Photonics.

• Janacek, G. (2010). Time series analysis forecasting and control. Journal of Time Series Analysis, 31(4), 303-303.

• Kalirajan, K. (1981). An econometric analysis of yield variability in paddy production. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, 29(3), 283-294.

• Katz, R. W., & Parlange, M. B. (1995). Generalizations of Chain-Dependent Processes: Application to Hourly Precipitation. Water Resources Research, 31(5), 1331-1341.

• Koucher, A. (1999). State and society: the question of agrarian change in Iraq 1921-1991.

• Lau, K. M., & Weng, H. (1995). Climate signal detection using wavelet transform: How to make a time series sing. Bulletin of the American Meteorological Society, 76(12), 2391-2402.

• Mamtimin, B., Et-Tantawi, A. M. M., Schaefer, D., Meixner, F. X., & Domroes, M. (2011). Recent trends of temperature change under hot and cold desert climates: Comparing the Sahara (Libya) and Central Asia (Xinjiang, China). Journal of Arid Environments, 75(11), 1105-1113.

• Mayer, J., Chambers, B., & Farooq, A. (Eds.). (1999). Development Policies in Natural Resource Economies. Edward Elgar Publishing.

• Monfreda, C., Ramankutty, N., & Foley, J. A. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global biogeochemical cycles, 22(1).

• Natrella, M. (2010). NIST/SEMATECH e-handbook of statistical methods.

• Ockerman, H. W., & Samano, S. G. (1985). The agricultural development of Iraq. Agricultural development in the Middle East., 189-207.

• Omer, T. M. A. (2011). Country pasture/forage resource profiles.

• Parry, M. L., Rosenzweig, C., Iglesias, A., Livermore, M., & Fischer, G. (2004). Effects of climate change on global food production under SRES emissions and socio-economic scenarios. Global Environmental Change, 14(1), 53-67.

• Pisias, N. G., & Moore, T. C. (1981). The evolution of Pleistocene climate: a time series approach. Earth and Planetary Science Letters, 52(2), 450-458.

• Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. PloS one, 8(6), e66428.

• Richardson, C. W. (1981). Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resources Research, 17(1), 182-190.