Fundamental statistical concepts

ANALYTICS PRIMER

3 Main Pieces of Statistics

Statistics and analytics in general boils down to three main pieces:

1. Data Collection
2. Data Analysis
3. Inference

Together these pieces summarize the data lifecycle from beginning to end.

3 Main Pieces of Statistics

Statistics and analytics in general boils down to three main pieces:

1. **Data Collection**
2. Data Analysis
3. Inference

One of the most overlooked pieces, but the most important!

- Bad data → Bad Results!

Populations vs. Samples

Population

- Set of all objects/individuals of interest
- Usually too large to obtain information from entire population

Sample

- Subset of the population that information is actually obtained

Populations vs. Samples

Population

- Set of all objects/individuals of interest
- Usually too large to obtain information from entire population

Sample

- Subset of the population that information is actually obtained
- **Sampling frame** – actual list from which the sample is taken

Populations vs. Samples

Population

- Set of all objects/individuals of interest
- Usually too large to obtain information from entire population

Sample

- Subset of the population that information is actually obtained
- **Sampling frame** – actual list from which the sample is taken → MAY NOT EQUAL POPULATION

Parameters vs. Statistics

Parameter

- Measures computed from a population.

Statistic

- Measures computed from a sample.
- Sample statistics is the **point estimate** of the population parameter.

Parameters vs. Statistics

Population	Sample
Parameter	Statistic

Parameters vs. Statistics

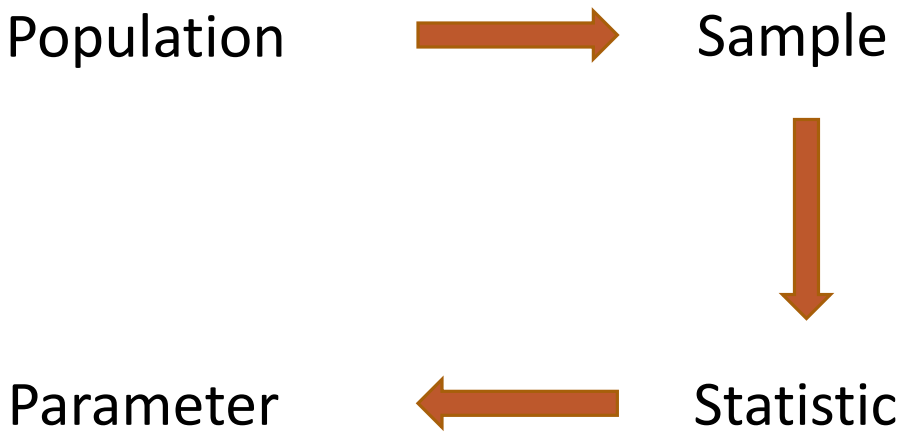
Population → Sample

Parameter

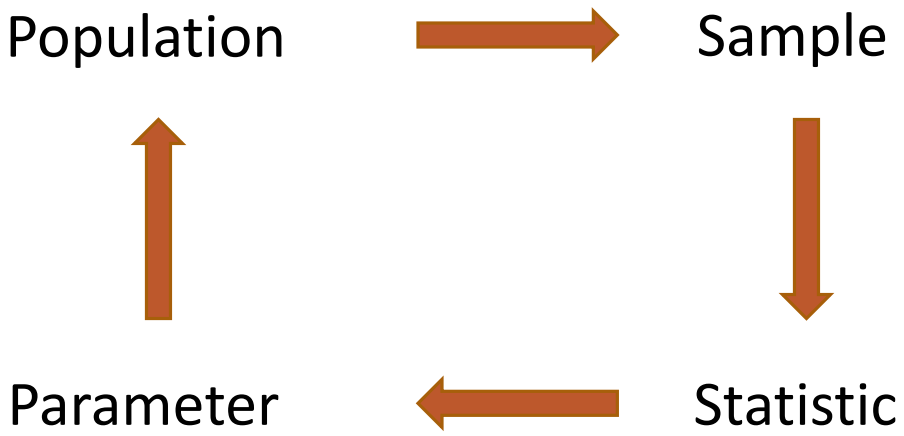
Statistic

Parameters vs. Statistics

Parameters vs. Statistics



Parameters vs. Statistics



Example

A retail chain is trying to determine if a new product they introduced is selling well across their stores. The retail chain has 2135 stores nationwide. The analyst in charge of this project is tasked to estimate the average daily sales of this new product across all stores. Older computing technology forces the company to randomly pick 179 stores spread evenly throughout the nation to gather the data. The average daily sales from these 179 stores is $129.19.

Identify population, sample, parameter, statistic.

Any sampling frame issues?

Example

A retail chain is trying to determine if a new product they introduced is selling well across their stores. The retail chain has **2135 stores nationwide**. The analyst in charge of this project is tasked to estimate the average daily sales of this new product across all stores. Older computing technology forces the company to randomly pick 179 stores spread evenly throughout the nation to calculate gather data from. The average daily sales from these 179 stores is $129.19.

Identify **population**, sample, parameter, statistic.

Any sampling frame issues?

Example

A retail chain is trying to determine if a new product they introduced is selling well across their stores. The retail chain has 2135 stores nationwide. The analyst in charge of this project is tasked to estimate the average daily sales of this new product across all stores. Older computing technology forces the company to randomly pick **179 stores spread evenly throughout the nation** to calculate gather data from. The average daily sales from these 179 stores is $129.19.

Identify population, **sample**, parameter, statistic.

Any sampling frame issues?

Example

A retail chain is trying to determine if a new product they introduced is selling well across their stores. The retail chain has 2135 stores nationwide. The analyst in charge of this project is tasked to estimate **the average daily sales of this new product across all stores**. Older computing technology forces the company to randomly pick 179 stores spread evenly throughout the nation to calculate gather data from. The average daily sales from these 179 stores is $129.19.

Identify population, sample, **parameter**, statistic.

Any sampling frame issues?

Example

A retail chain is trying to determine if a new product they introduced is selling well across their stores. The retail chain has 2135 stores nationwide. The analyst in charge of this project is tasked to estimate the average daily sales of this new product across all stores. Older computing technology forces the company to randomly pick 179 stores spread evenly throughout the nation to calculate gather data from. The average daily sales from these 179 stores is **$129.19**.

Identify population, sample, parameter, **statistic**.

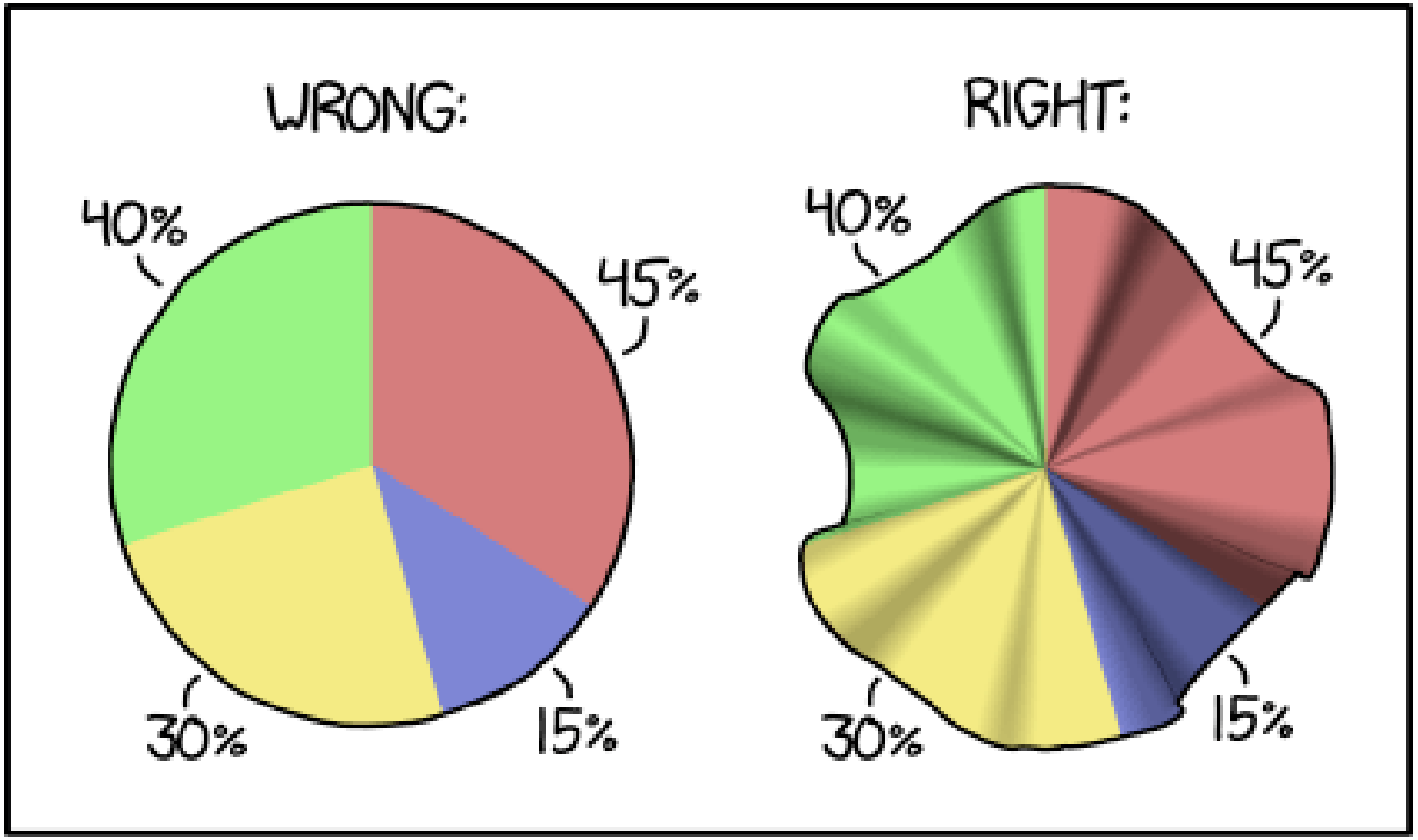
Any sampling frame issues?

Example

A retail chain is trying to determine if a new product they introduced is selling well across their stores. The retail chain has 2135 stores nationwide. The analyst in charge of this project is tasked to estimate the average daily sales of this new product across all stores. Older computing technology forces the company to randomly pick 179 stores spread evenly throughout the nation to calculate gather data from. The average daily sales from these 179 stores is $129.19.

Identify population, sample, parameter, statistic.

Any sampling frame issues? **NO**



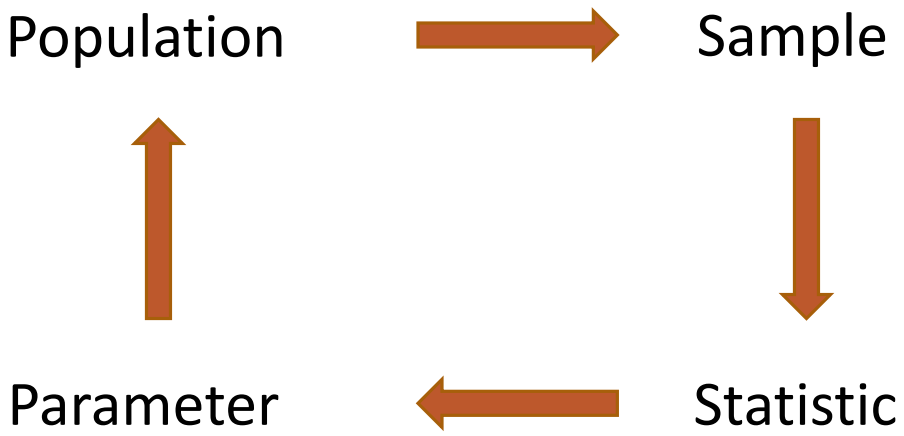
HOW TO MAKE A PIE CHART IF YOUR PERCENTAGES DON'T ADD UP TO 100

Source: xkcd.com

Sampling Techniques

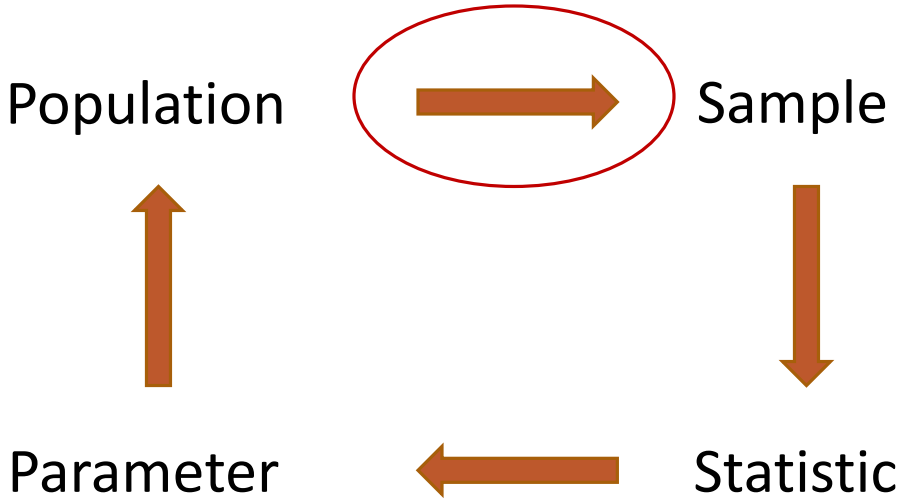
FUNDAMENTAL STATISTICAL CONCEPTS (2 OF 3)

Parameters vs. Statistics



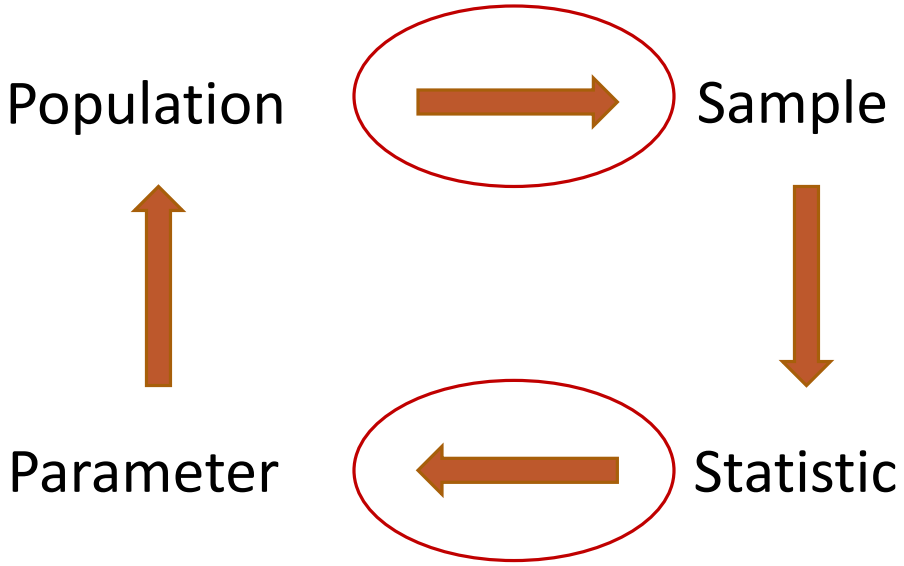
Parameters vs. Statistics

Need good sampling to…



Parameters vs. Statistics

Need good sampling to…



…have good estimates.

Sampling

There are many different ways to sample data from population.

Mistakes in sampling can lead to bias in the sample.

Bias – certain outcomes are favored over other outcomes in samples.

Types of Bias

Bias – certain outcomes are favored over other outcomes in samples.

2 Common Types of Bias:

1. Sampling Bias
2. Selection Bias

Types of Bias

Bias – certain outcomes are favored over other outcomes in samples.

2 Common Types of Bias:

1. Sampling Bias
2. Selection Bias

Types of Bias

Bias – certain outcomes are favored over other outcomes in samples.

2 Common Types of Bias:

1. Sampling Bias
   a) **Convenience sampling** – technique that selects subjects from population based on accessibility and ease.
   b) **Voluntary sampling** – technique where subjects volunteer themselves to sample.
2. Selection Bias

Types of Bias

Bias – certain outcomes are favored over other outcomes in samples.

2 Common Types of Bias:

1. Sampling Bias
2. Selection Bias

Types of Bias

Bias – certain outcomes are favored over other outcomes in samples.

2 Common Types of Bias:

1. Sampling Bias
2. Selection Bias
   a) **Undercoverage** – frame and population are not equal (ex. Phone book)
   b) **Nonresponse** – subject in sample cannot / will not respond or be measured (ex. Telemarketer)

Statistical Techniques

Statistical sampling techniques use selection methods based on chance selection instead of convenience or judgement.

4 Common Techniques:

1. Simple Random Sampling (SRS)
2. Stratified Random Sampling
3. Cluster Sampling
4. Systematic Sampling

Simple Random Sampling (SRS)

A method of sampling items from a population such that every possible sample of a specified size has an equal chance of being selected.

Sampling Frame

#BringBackOurGirls #BringBackOurGirls #BringBackOurGirls #BringBackOurGirls

1. Blessing Abana
2. Yayi Abana
3. Deborah Abari
4. Deborah Abbas
5. Hauwa Abdu
6. Ihyi Abdu
7. Safiya Abdu
8. Sieker* Abdul
9. Awa Abge
10. Deborah Abge
11. Kummai Aboku
12. Hamsatu Abubakar
13. Mairama Abubakar
14. Hasana Adamu
15. Naomi Adamu
16. Amina Ali
17. Asabe Ali
18. Mary Ali
19. Comfort Amos
20. Deborah Amos
21. Mary Amos*
22. Rifkatu Amos
23. Ruth Amos
24. Saraya Amos
25. Esther Ayuba
26. Saratu Ayuba
27. Hauwa Balti
28. Awa Bitrus
29. Christiana Bitrus
30. Godiya Bitrus
31. Na'omi Bitrus
32. Rahila Bitrus
33. Ruth Bitrus
34. Nguba Buba
35. Maryamu Bulama
36. Abigail Bukar
37. Yana Bukar
38. Gloria Dama
39. Maifa Dama
40. Muwa Daniel
41. Talata Daniel
42. Filo Dauda
43. Mary Dauda
44. Saratu Dauda
45. Patiant Dzakwa
46. Saratu Emmanuel
47. Lydia Emmar
48. Monica Enoch
49. Aisha Ezekial
50. Rifkatu Galang
51. Halima Gamba
52. Aishatu Grema
53. Comfort* Habila
54. Lydia Habila
55. Liyatu Habitu
56. Febi Haruna
57. Tabitha Hyelampa
58. Ladi Ibrahim
59. Rahap Ibrahim
60. Hanatu Ishaku
61. Ruth Ishaku
62. Zara Ishaku
63. Hauwa Ishaya
64. Hauwa Isuwa
65. Awa James
66. Deborah Jafaru
67. Ladi Joel
68. Esther John
69. Jummai John
70. Laraba John
71. Rhoda John
72. Eli Joseph
73. Esther Joshua
74. Hauwa Joseph
75. Rebeca Joseph
76. Ruth Joshua
77. Yana Joshua
78. Rebeca Kabu
79. Ruth Kollo
80. Hauwa Kwakwi
81. Rakiya Kwamtah
82. Kauna Lalai
83. Aisha Lawan
84. Falta Lawan
85. Maryamu Lawan
86. Ruth Lawan
87. Kauna Luka
88. Na'omi Luka
89. Rebecca Luka
90. Gloria Mainta
91. Kabu Malla
92. Rebecca Mallum
93. Laraba Maman
94. Asabe Manu
95. Kwadugu Manu
96. Esther Markus
97. Saratu Markus
98. Aishatu Musa
99. Awagana Musa
100. Hanatu Musa
101. Hauwa Musa
102. Helen Musa
103. Maryamu Musa
104. Palmata Musa
105. Rejoice Musa
106. Saraya Musa
107. Hauwa Mutah
108. Kume Mutah
109. Lugwa Mutah
110. Ruth Ngladar
111. Hauwa Nkeki
112. Racheal Nkeki
113. Hanatu Nuhu
114. Pindar Nuhu
115. Grace Paul
116. Jummai Paul
117. Laraba Paul
118. Ladi Paul
119. Mary Paul
120. Saraya Paul
121. Deborah Peter
122. Hauwa Peter
123. Rhoda Peter
124. Naomi Philimon
125. Mwa Malam Pogu
126. Salomi Pogu
127. Tabitha Pogu
128. Yana Pogu
129. Luggwa Samuel
130. Saraya Samuel
131. Serah Samuel
132. Luggwa Sanda
133. Rejoice Sanki
134. Margret Shettima
135. Yanke Shittima
136. Tabitha Silas
137. Kwanta Simon
138. Lydia Simon
139. Deborah Soloman
140. Rifkatu Soloman
141. Hana Stephen
142. Mary Sule
143. Saraya Mal Stover
144. Fatima Tabji
145. Hauwa Takai
146. Hauwa Tella
147. Solomi Titus
148. Esther Usman
149. Maimuna Usman
150. Mary Usman
151. Ladi Wadai
152. Muli Waligam
153. Margret Watsai
154. Maryamu Wavi
155. Hauwa Wule
156. Glory Yaga
157. Mairama Yahaya
158. Christy Yahi
159. Anthonia Yahonna
160. Debora Yahonna
161. Laraba Yahonna
162. Mary Yahonna*
163. Na'omi Yahonna
164. Rahila Yahanna
165. Dorcas* Yakubu
166. Hadiza Yakubu
167. Juliana Yakubu
168. Mary Yakubu
169. Maryamu Yakubu
170. Rifkatu Yakubu
171. Suzana Yakubu
172. Jinkai Yama
173. Margret Yama
174. Saraya Yanga
175. Yana Yidau
176. Hauwa Yirma
177. Naomi Zakaria

* = Spelling unconfirmed; names respectfully corrected to most likely spelling. Retyped by @ElisabethEpps.

15 April, 2014: At least 276 girls aged 15-18 yo were kidnapped from their boarding school in Chibok, Nigeria.
4 May, 2014: Christian Association of Nigeria released this list of 177 names reportedly confirmed still missing.

Simple Random Sampling (SRS)

A method of sampling items from a population such that every possible sample of a specified size has an equal chance of being selected.

Sampling Frame

#BringBackOurGirls #BringBackOurGirls #BringBackOurGirls #BringBackOurGirls

1. Blessing Abana
2. Yayi Abana
3. Deborah Abari
4. Deborah Abbas
5. Hauwa Abdu
6. Ihyi Abdu
7. Safiya Abdu
8. Sieker* Abdul
9. Awa Abge
10. Deborah Abge
11. Kummai Aboku
12. Hamsatu Abubakar
13. Mairama Abubakar
14. Hasana Adamu
15. Naomi Adamu
16. Amina Ali
17. Asabe Ali
18. Mary Ali
19. Comfort Amos
20. Deborah Amos
21. Mary Amos*
22. Rifkatu Amos
23. Ruth Amos
24. Saraya Amos
25. Esther Ayuba
26. Saratu Ayuba
27. Hauwa Balti
28. Awa Bitrus
29. Christiana Bitrus
30. Godiya Bitrus
31. Na'omi Bitrus
32. Rahila Bitrus
33. Ruth Bitrus
34. Nguba Buba
35. Maryamu Bulama
36. Abigail Bukar
37. Yana Bukar
38. Gloria Dama
39. Maifa Dama
40. Muwa Daniel
41. Talata Daniel
42. Filo Dauda
43. Mary Dauda
44. Saratu Dauda
45. Patiant Dzakwa
46. Saratu Emmanuel
47. Lydia Emmar
48. Monica Enoch
49. Aisha Ezekial
50. Rifkatu Galang
51. Halima Gamba
52. Aishatu Grema
53. Comfort* Habila
54. Lydia Habila
55. Liyatu Habitu
56. Febi Haruna
57. Tabitha Hyelampa
58. Ladi Ibrahim
59. Rahap Ibrahim
60. Hanatu Ishaku
61. Ruth Ishaku
62. Zara Ishaku
63. Hauwa Ishaya
64. Hauwa Isuwa
65. Awa James
66. Deborah Jafaru
67. Ladi Joel
68. Esther John
69. Jummai John
70. Laraba John
71. Rhoda John
72. Eli Joseph
73. Esther Joshua
74. Hauwa Joseph
75. Rebeca Joseph
76. Ruth Joshua
77. Yana Joshua
78. Rebeca Kabu
79. Ruth Kollo
80. Hauwa Kwakwi
81. Rakiya Kwamtah
82. Kauna Lalai
83. Aisha Lawan
84. Falta Lawan
85. Maryamu Lawan
86. Ruth Lawan
87. Kauna Luka
88. Na'omi Luka
89. Rebecca Luka
90. Gloria Mainta
91. Kabu Malla
92. Rebecca Mallum
93. Laraba Maman
94. Asabe Manu
95. Kwadugu Manu
96. Esther Markus
97. Saratu Markus
98. Aishatu Musa
99. Awagana Musa
100. Hanatu Musa
101. Hauwa Musa
102. Helen Musa
103. Maryamu Musa
104. Palmata Musa
105. Rejoice Musa
106. Saraya Musa
107. Hauwa Mutah
108. Kume Mutah
109. Lugwa Mutah
110. Ruth Ngladar
111. Hauwa Nkeki
112. Racheal Nkeki
113. Hanatu Nuhu
114. Pindar Nuhu
115. Grace Paul
116. Jummai Paul
117. Laraba Paul
118. Ladi Paul
119. Mary Paul
120. Saraya Paul
121. Deborah Peter
122. Hauwa Peter
123. Rhoda Peter
124. Naomi Philimon
125. Mwa Malam Pogu
126. Salomi Pogu
127. Tabitha Pogu
128. Yana Pogu
129. Luggwa Samuel
130. Saraya Samuel
131. Serah Samuel
132. Luggwa Sanda
133. Rejoice Sanki
134. Margret Shettima
135. Yanke Shittima
136. Tabitha Silas
137. Kwanta Simon
138. Lydia Simon
139. Deborah Soloman
140. Rifkatu Soloman
141. Hana Stephen
142. Mary Sule
143. Saraya Mal Stover
144. Fatima Tabji
145. Hauwa Takai
146. Hauwa Tella
147. Solomi Titus
148. Esther Usman
149. Maimuna Usman
150. Mary Usman
151. Ladi Wadai
152. Mali Waligam
153. Margret Watsai
154. Maryamu Wavi
155. Hauwa Wule
156. Glory Yaga
157. Mairama Yahaya
158. Christy Yahi
159. Anthonia Yahonna
160. Debora Yahonna
161. Laraba Yahonna
162. Mary Yahonna*
163. Na'omi Yahonna
164. Rahila Yahanna
165. Dorcas* Yakubu
166. Hadiza Yakubu
167. Juliana Yakubu
168. Mary Yakubu
169. Maryamu Yakubu
170. Rifkatu Yakubu
171. Suzana Yakubu
172. Jinkai Yama
173. Margret Yama
174. Saraya Yanga
175. Yana Yidau
176. Hauwa Yirma
177. Naomi Zakaria

* = Spelling unconfirmed; names respectfully corrected to most likely spelling. Retyped by @ElisabethEpps.

15 April, 2014: At least 276 girls aged 15-18 yo were kidnapped from their boarding school in Chibok, Nigeria.
4 May, 2014: Christian Association of Nigeria released this list of 177 names reportedly confirmed still missing.

Simple Random Sampling (SRS)

A method of sampling items from a population such that every possible sample of a specified size has an equal chance of being selected.

Advantages:

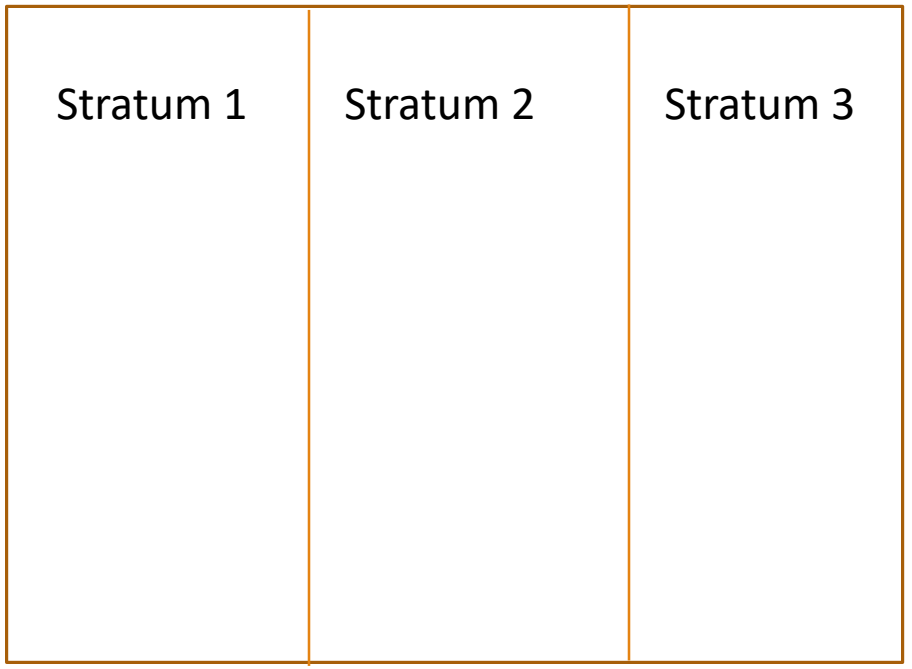
- No statistical bias, no previous information about sample needed ahead of time

Disadvantages:

- Expensive, time consuming, hard to implement (need list of population)

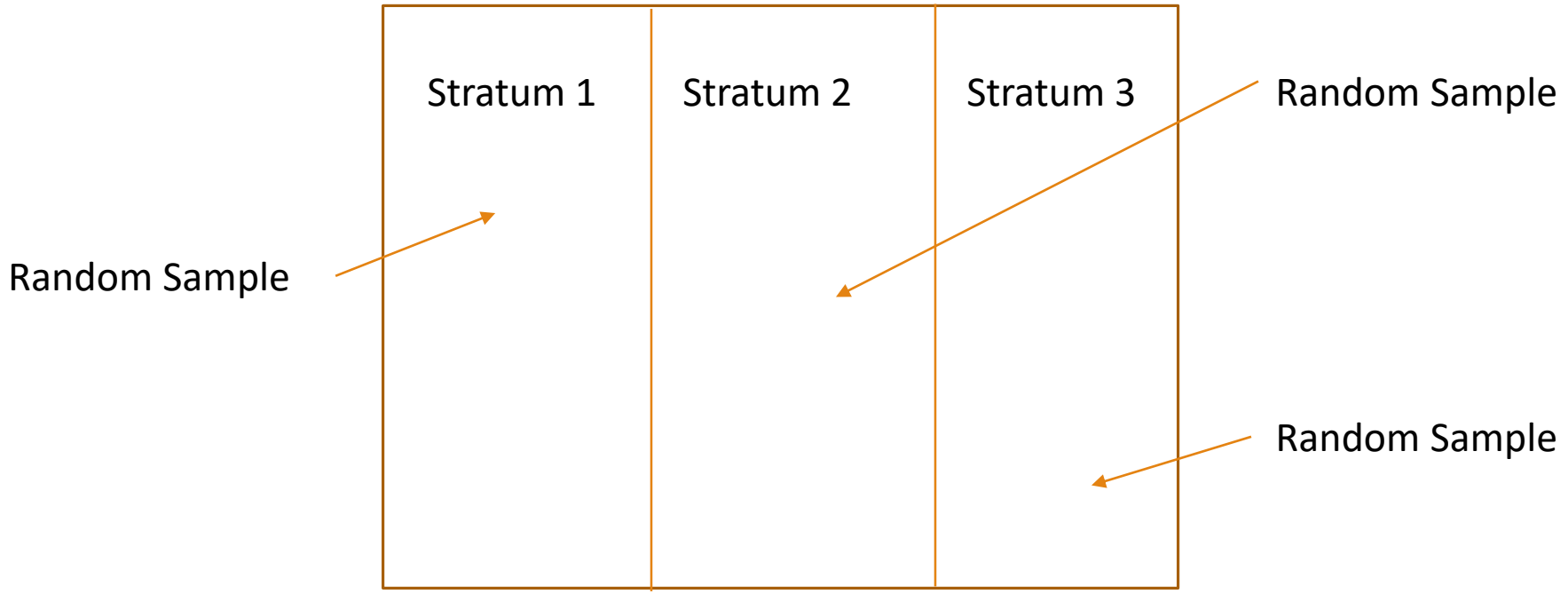
Stratified Random Sampling (STS)

A method of sampling items where the population is divided *a priori* into subgroups, called **strata**, so that each member in the population belongs to only one stratum. Sample items from **every** stratum (with SRS for example).



Stratified Random Sampling (STS)

A method of sampling items where the population is divided *a priori* into subgroups, called **strata**, so that each member in the population belongs to only one stratum. Sample items from **every** stratum (with SRS for example).



Stratified Random Sampling (STS)

A method of sampling items where the population is divided *a priori* into subgroups, called **strata**, so that each member in the population belongs to only one strata. Sample items from **every** strata (with SRS for example).

Advantages:

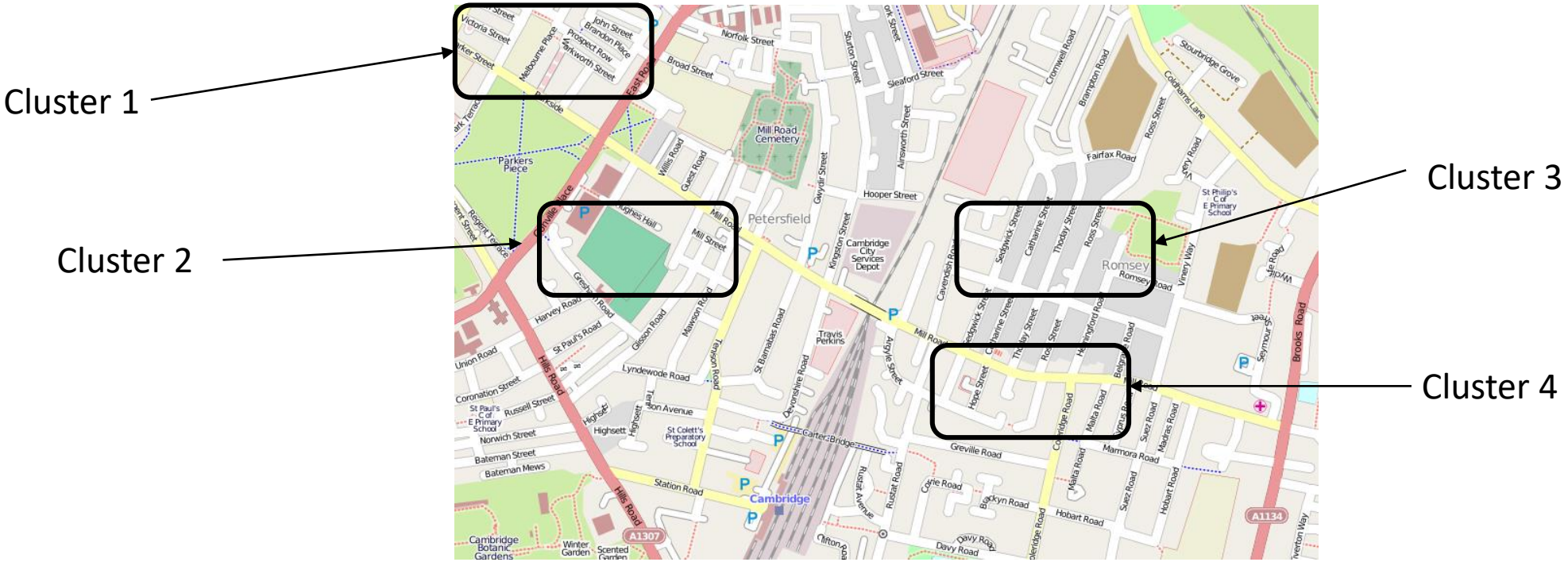
- Smaller sample sizes can achieve same accuracy as SRS, more information about parts of population

Disadvantages:

- Need information about population ahead of time to split on!

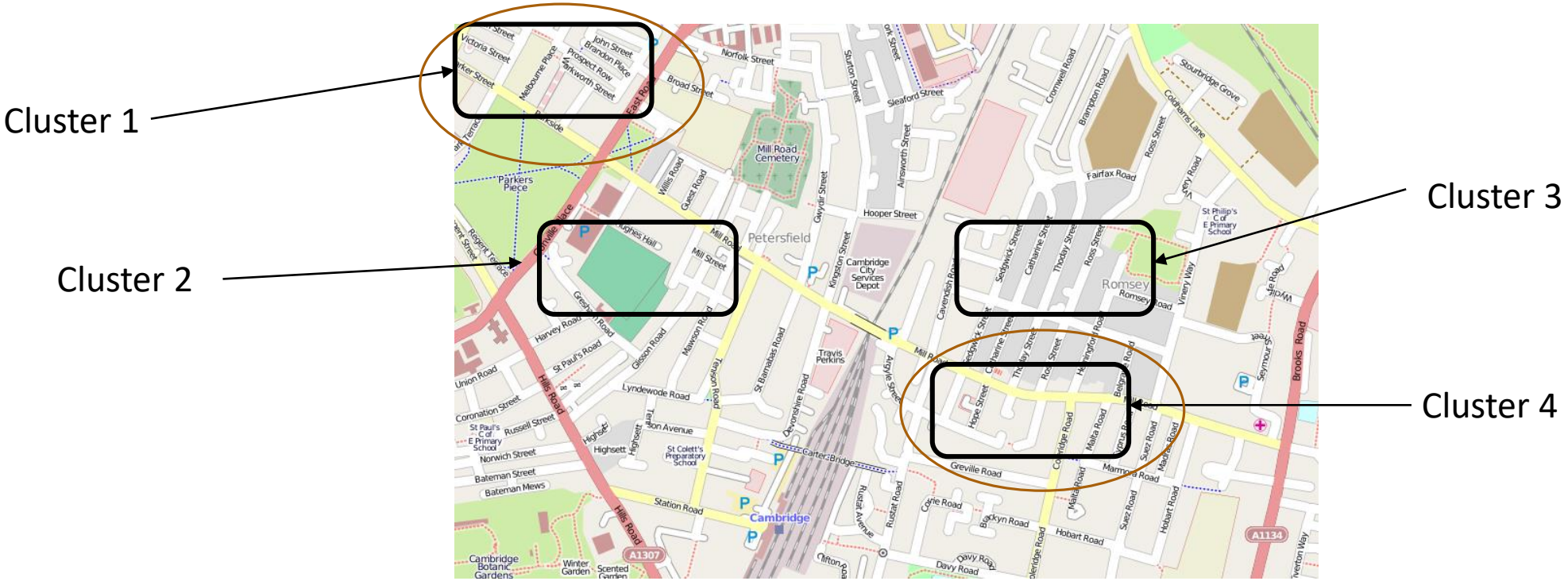
Cluster Sampling

A method of sampling items where the population is divided *a priori* into subgroups, called **clusters**, so that each member in the population belongs to only one cluster. Sample items from **a sample** of *m* clusters selected randomly.



Cluster Sampling

A method of sampling items where the population is divided *a priori* into subgroups, called **clusters**, so that each member in the population belongs to only one cluster. Sample items from **a sample** of *m* clusters selected randomly.



Cluster Sampling

A method of sampling items where the population is divided *a priori* into subgroups, called **clusters**, so that each member in the population belongs to only one cluster. Sample items from **a sample** of *m* clusters selected randomly.

Advantages:

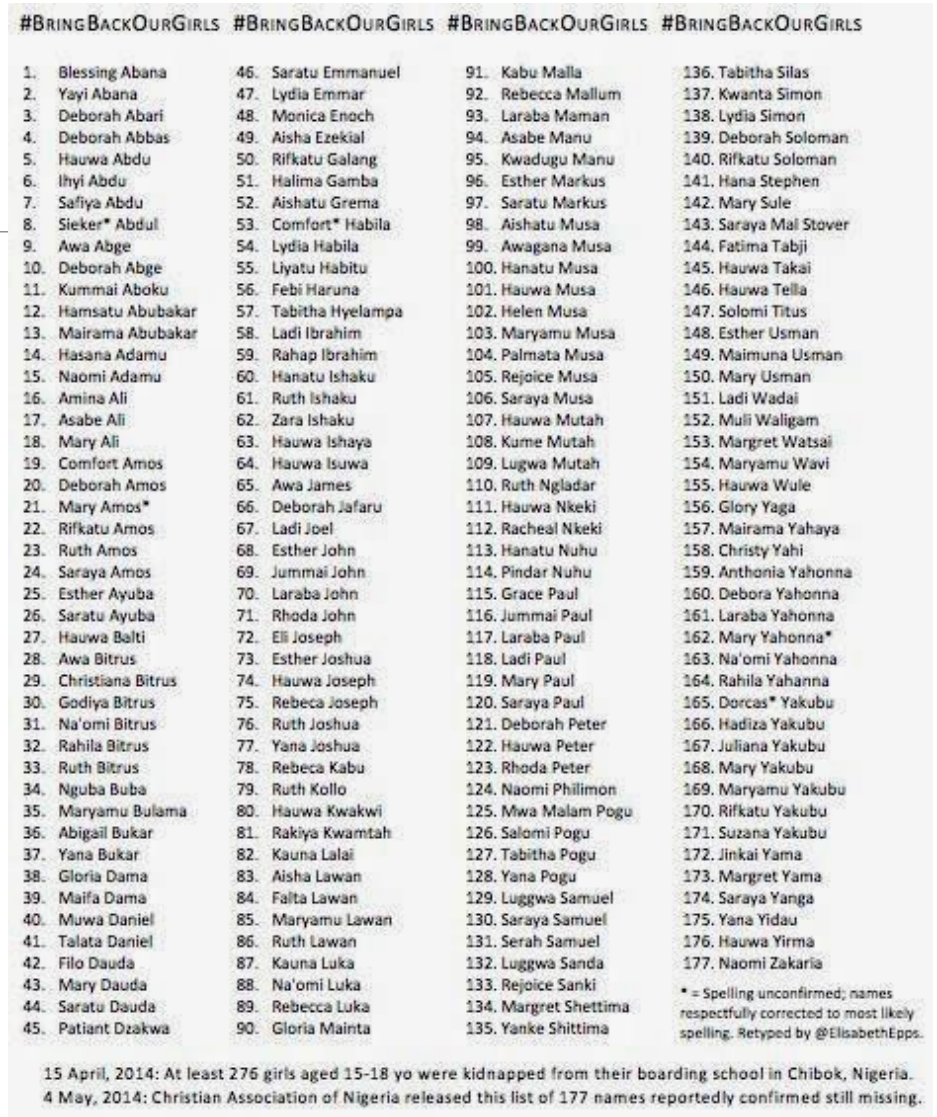
- Overcome issues with travel, time, and expense; Easier to implement than SRS or STS

Disadvantages:

- May have slight bias if random clusters aren't representative; actual sample size depends on number of units within each cluster sampled

Systematic Sampling

A method of sampling items that involves selecting every k^{th} item in the population after randomly selecting a starting point between 1 and k. The value k is determined as the ratio of the population size over the desired sample size.



#BringBackOurGirls #BringBackOurGirls #BringBackOurGirls #BringBackOurGirls

1. Blessing Abana
2. Yayi Abana
3. Deborah Abari
4. Deborah Abbas
5. Hauwa Abdu
6. Ihyi Abdu
7. Safiya Abdu
8. Sieker* Abdul
9. Awa Abge
10. Deborah Abge
11. Kummai Aboku
12. Hamsatu Abubakar
13. Mairama Abubakar
14. Hasana Adamu
15. Naomi Adamu
16. Amina Ali
17. Asabe Ali
18. Mary Ali
19. Comfort Amos
20. Deborah Amos
21. Mary Amos*
22. Rifkatu Amos
23. Ruth Amos
24. Saraya Amos
25. Esther Ayuba
26. Saratu Ayuba
27. Hauwa Balti
28. Awa Bitrus
29. Christiana Bitrus
30. Godiya Bitrus
31. Na'omi Bitrus
32. Rahila Bitrus
33. Ruth Bitrus
34. Nguba Buba
35. Maryamu Bulama
36. Abigail Bukar
37. Yana Bukar
38. Gloria Dama
39. Maifa Dama
40. Muwa Daniel
41. Talata Daniel
42. Filo Dauda
43. Mary Dauda
44. Saratu Dauda
45. Patiant Dzakwa
46. Saratu Emmanuel
47. Lydia Emmar
48. Monica Enoch
49. Aisha Ezekial
50. Rifkatu Galang
51. Halima Gamba
52. Aishatu Grema
53. Comfort* Habila
54. Lydia Habila
55. Liyatu Habitu
56. Febi Haruna
57. Tabitha Hyelampa
58. Ladi Ibrahim
59. Rahap Ibrahim
60. Hanatu Ishaku
61. Ruth Ishaku
62. Zara Ishaku
63. Hauwa Ishaya
64. Hauwa Isuwa
65. Awa James
66. Deborah Jafaru
67. Ladi Joel
68. Esther John
69. Jummai John
70. Laraba John
71. Rhoda John
72. Eli Joseph
73. Esther Joshua
74. Hauwa Joseph
75. Rebeca Joseph
76. Ruth Joshua
77. Yana Joshua
78. Rebeca Kabu
79. Ruth Kollo
80. Hauwa Kwakwi
81. Rakiya Kwamtah
82. Kauna Lalai
83. Aisha Lawan
84. Falta Lawan
85. Maryamu Lawan
86. Ruth Lawan
87. Kauna Luka
88. Na'omi Luka
89. Rebecca Luka
90. Gloria Mainta
91. Kabu Malla
92. Rebecca Mallum
93. Laraba Maman
94. Asabe Manu
95. Kwadugu Manu
96. Esther Markus
97. Saratu Markus
98. Aishatu Musa
99. Awagana Musa
100. Hanatu Musa
101. Hauwa Musa
102. Helen Musa
103. Maryamu Musa
104. Palmata Musa
105. Rejoice Musa
106. Saraya Musa
107. Hauwa Mutah
108. Kume Mutah
109. Lugwa Mutah
110. Ruth Ngladar
111. Hauwa Nkeki
112. Racheal Nkeki
113. Hanatu Nuhu
114. Pindar Nuhu
115. Grace Paul
116. Jummai Paul
117. Laraba Paul
118. Ladi Paul
119. Mary Paul
120. Saraya Paul
121. Deborah Peter
122. Hauwa Peter
123. Rhoda Peter
124. Naomi Philimon
125. Mwa Malam Pogu
126. Salomi Pogu
127. Tabitha Pogu
128. Yana Pogu
129. Luggwa Samuel
130. Saraya Samuel
131. Serah Samuel
132. Luggwa Sanda
133. Rejoice Sanki
134. Margret Shettima
135. Yanke Shittima
136. Tabitha Silas
137. Kwanta Simon
138. Lydia Simon
139. Deborah Soloman
140. Rifkatu Soloman
141. Hana Stephen
142. Mary Sule
143. Saraya Mal Stover
144. Fatima Tabji
145. Hauwa Takai
146. Hauwa Tella
147. Solomi Titus
148. Esther Usman
149. Maimuna Usman
150. Mary Usman
151. Ladi Wadai
152. Muli Waligam
153. Margret Watsai
154. Maryamu Wavi
155. Hauwa Wule
156. Glory Yaga
157. Mairama Yahaya
158. Christy Yahi
159. Anthonia Yahonna
160. Debora Yahonna
161. Laraba Yahonna
162. Mary Yahonna*
163. Na'omi Yahonna
164. Rahila Yahanna
165. Dorcas* Yakubu
166. Hadiza Yakubu
167. Juliana Yakubu
168. Mary Yakubu
169. Maryamu Yakubu
170. Rifkatu Yakubu
171. Suzana Yakubu
172. Jinkai Yama
173. Margret Yama
174. Saraya Yanga
175. Yana Yidau
176. Hauwa Yirma
177. Naomi Zakaria

* = Spelling unconfirmed; names respectfully corrected to most likely spelling. Retyped by @ElisabethEpps.

15 April, 2014: At least 276 girls aged 15-18 yo were kidnapped from their boarding school in Chibok, Nigeria.
4 May, 2014: Christian Association of Nigeria released this list of 177 names reportedly confirmed still missing.

Systematic Sampling

A method of sampling items that involves selecting every k^{th} item in the population after randomly selecting a starting point between 1 and k. The value k is determined as the ratio of the population size over the desired sample size.

Random Starting point

Say k=5

#BringBackOurGirls #BringBackOurGirls #BringBackOurGirls #BringBackOurGirls

1. Blessing Abana
2. Yayi Abana
3. Deborah Abari
4. Deborah Abbas
5. Hauwa Abdu
6. Ihyi Abdu
7. Safiya Abdu
8. Sieker* Abdul
9. Awa Abge
10. Deborah Abge
11. Kummai Aboku
12. Hamsatu Abubakar
13. Mairama Abubakar
14. Hasana Adamu
15. Naomi Adamu
16. Amina Ali
17. Asabe Ali
18. Mary Ali
19. Comfort Amos
20. Deborah Amos
21. Mary Amos*
22. Rifkatu Amos
23. Ruth Amos
24. Saraya Amos
25. Esther Ayuba
26. Saratu Ayuba
27. Hauwa Balti
28. Awa Bitrus
29. Christiana Bitrus
30. Godiya Bitrus
31. Na'omi Bitrus
32. Rahila Bitrus
33. Ruth Bitrus
34. Nguba Buba
35. Maryamu Bulama
36. Abigail Bukar
37. Yana Bukar
38. Gloria Dama
39. Maifa Dama
40. Muwa Daniel
41. Talata Daniel
42. Filo Dauda
43. Mary Dauda
44. Saratu Dauda
45. Patiant Dzakwa
46. Saratu Emmanuel
47. Lydia Emmar
48. Monica Enoch
49. Aisha Ezekial
50. Rifkatu Galang
51. Halima Gamba
52. Aishatu Grema
53. Comfort* Habila
54. Lydia Habila
55. Liyatu Habitu
56. Febi Haruna
57. Tabitha Hyelampa
58. Ladi Ibrahim
59. Rahap Ibrahim
60. Hanatu Ishaku
61. Ruth Ishaku
62. Zara Ishaku
63. Hauwa Ishaya
64. Hauwa Isuwa
65. Awa James
66. Deborah Jafaru
67. Ladi Joel
68. Esther John
69. Jummai John
70. Laraba John
71. Rhoda John
72. Eli Joseph
73. Esther Joshua
74. Hauwa Joseph
75. Rebeca Joseph
76. Ruth Joshua
77. Yana Joshua
78. Rebeca Kabu
79. Ruth Kollo
80. Hauwa Kwakwi
81. Rakiya Kwamtah
82. Kauna Lalai
83. Aisha Lawan
84. Falta Lawan
85. Maryamu Lawan
86. Ruth Lawan
87. Kauna Luka
88. Na'omi Luka
89. Rebecca Luka
90. Gloria Mainta
91. Kabu Malla
92. Rebecca Mallum
93. Laraba Maman
94. Asabe Manu
95. Kwadugu Manu
96. Esther Markus
97. Saratu Markus
98. Aishatu Musa
99. Awagana Musa
100. Hanatu Musa
101. Hauwa Musa
102. Helen Musa
103. Maryamu Musa
104. Palmata Musa
105. Rejoice Musa
106. Saraya Musa
107. Hauwa Mutah
108. Kume Mutah
109. Lugwa Mutah
110. Ruth Ngladar
111. Hauwa Nkeki
112. Racheal Nkeki
113. Hanatu Nuhu
114. Pindar Nuhu
115. Grace Paul
116. Jummai Paul
117. Laraba Paul
118. Ladi Paul
119. Mary Paul
120. Saraya Paul
121. Deborah Peter
122. Hauwa Peter
123. Rhoda Peter
124. Naomi Philimon
125. Mwa Malam Pogu
126. Salomi Pogu
127. Tabitha Pogu
128. Yana Pogu
129. Luggwa Samuel
130. Saraya Samuel
131. Serah Samuel
132. Luggwa Sanda
133. Rejoice Sanki
134. Margret Shettima
135. Yanke Shittima
136. Tabitha Silas
137. Kwanta Simon
138. Lydia Simon
139. Deborah Soloman
140. Rifkatu Soloman
141. Hana Stephen
142. Mary Sule
143. Saraya Mal Stover
144. Fatima Tabji
145. Hauwa Takai
146. Hauwa Tella
147. Solomi Titus
148. Esther Usman
149. Maimuna Usman
150. Mary Usman
151. Ladi Wadai
152. Muli Waligam
153. Margret Watsai
154. Maryamu Wavi
155. Hauwa Wule
156. Glory Yaga
157. Mairama Yahaya
158. Christy Yahi
159. Anthonia Yahonna
160. Debora Yahonna
161. Laraba Yahonna
162. Mary Yahonna*
163. Na'omi Yahonna
164. Rahila Yahanna
165. Dorcas* Yakubu
166. Hadiza Yakubu
167. Juliana Yakubu
168. Mary Yakubu
169. Maryamu Yakubu
170. Rifkatu Yakubu
171. Suzana Yakubu
172. Jinkai Yama
173. Margret Yama
174. Saraya Yanga
175. Yana Yidau
176. Hauwa Yirma
177. Naomi Zakaria

* = Spelling unconfirmed; names respectfully corrected to most likely spelling. Retyped by @ElisabethEpps.

15 April, 2014: At least 276 girls aged 15-18 yo were kidnapped from their boarding school in Chibok, Nigeria.
4 May, 2014: Christian Association of Nigeria released this list of 177 names reportedly confirmed still missing.

Systematic Sampling

A method of sampling items that involves selecting every k^{th} item in the population after randomly selecting a starting point between 1 and k. The value k is determined as the ratio of the population size over the desired sample size.

Random Starting point

And so on….til you reach desired sample

#BringBackOurGirls #BringBackOurGirls #BringBackOurGirls #BringBackOurGirls

1. Blessing Abana
2. Yayi Abana
3. Deborah Abari
4. Deborah Abbas
5. Hauwa Abdu
6. Ihyi Abdu
7. Safiya Abdu
8. Sieker* Abdul
9. Awa Abge
10. Deborah Abge
11. Kummai Aboku
12. Hamsatu Abubakar
13. Mairama Abubakar
14. Hasana Adamu
15. Naomi Adamu
16. Amina Ali
17. Asabe Ali
18. Mary Ali
19. Comfort Amos
20. Deborah Amos
21. Mary Amos*
22. Rifkatu Amos
23. Ruth Amos
24. Saraya Amos
25. Esther Ayuba
26. Saratu Ayuba
27. Hauwa Balti
28. Awa Bitrus
29. Christiana Bitrus
30. Godiya Bitrus
31. Na'omi Bitrus
32. Rahila Bitrus
33. Ruth Bitrus
34. Nguba Buba
35. Maryamu Bulama
36. Abigail Bukar
37. Yana Bukar
38. Gloria Dama
39. Maifa Dama
40. Muwa Daniel
41. Talata Daniel
42. Filo Dauda
43. Mary Dauda
44. Saratu Dauda
45. Patiant Dzakwa
46. Saratu Emmanuel
47. Lydia Emmar
48. Monica Enoch
49. Aisha Ezekial
50. Rifkatu Galang
51. Halima Gamba
52. Aishatu Grema
53. Comfort* Habila
54. Lydia Habila
55. Liyatu Habitu
56. Febi Haruna
57. Tabitha Hyelampa
58. Ladi Ibrahim
59. Rahap Ibrahim
60. Hanatu Ishaku
61. Ruth Ishaku
62. Zara Ishaku
63. Hauwa Ishaya
64. Hauwa Isuwa
65. Awa James
66. Deborah Jafaru
67. Ladi Joel
68. Esther John
69. Jummai John
70. Laraba John
71. Rhoda John
72. Eli Joseph
73. Esther Joshua
74. Hauwa Joseph
75. Rebeca Joseph
76. Ruth Joshua
77. Yana Joshua
78. Rebeca Kabu
79. Ruth Kollo
80. Hauwa Kwakwi
81. Rakiya Kwamtah
82. Kauna Lalai
83. Aisha Lawan
84. Falta Lawan
85. Maryamu Lawan
86. Ruth Lawan
87. Kauna Luka
88. Na'omi Luka
89. Rebecca Luka
90. Gloria Mainta
91. Kabu Malla
92. Rebecca Mallum
93. Laraba Maman
94. Asabe Manu
95. Kwadugu Manu
96. Esther Markus
97. Saratu Markus
98. Aishatu Musa
99. Awagana Musa
100. Hanatu Musa
101. Hauwa Musa
102. Helen Musa
103. Maryamu Musa
104. Palmata Musa
105. Rejoice Musa
106. Saraya Musa
107. Hauwa Mutah
108. Kume Mutah
109. Lugwa Mutah
110. Ruth Ngladar
111. Hauwa Nkeki
112. Racheal Nkeki
113. Hanatu Nuhu
114. Pindar Nuhu
115. Grace Paul
116. Jummai Paul
117. Laraba Paul
118. Ladi Paul
119. Mary Paul
120. Saraya Paul
121. Deborah Peter
122. Hauwa Peter
123. Rhoda Peter
124. Naomi Philimon
125. Mwa Malam Pogu
126. Salomi Pogu
127. Tabitha Pogu
128. Yana Pogu
129. Luggwa Samuel
130. Saraya Samuel
131. Serah Samuel
132. Luggwa Sanda
133. Rejoice Sanki
134. Margret Shettima
135. Yanke Shittima
136. Tabitha Silas
137. Kwanta Simon
138. Lydia Simon
139. Deborah Soloman
140. Rifkatu Soloman
141. Hana Stephen
142. Mary Sule
143. Saraya Mal Stover
144. Fatima Tabji
145. Hauwa Takai
146. Hauwa Tella
147. Solomi Titus
148. Esther Usman
149. Maimuna Usman
150. Mary Usman
151. Ladi Wadai
152. Muli Waligam
153. Margret Watsai
154. Maryamu Wavi
155. Hauwa Wule
156. Glory Yaga
157. Mairama Yahaya
158. Christy Yahi
159. Anthonia Yahonna
160. Debora Yahonna
161. Laraba Yahonna
162. Mary Yahonna*
163. Na'omi Yahonna
164. Rahila Yahanna
165. Dorcas* Yakubu
166. Hadiza Yakubu
167. Juliana Yakubu
168. Mary Yakubu
169. Maryamu Yakubu
170. Rifkatu Yakubu
171. Suzana Yakubu
172. Jinkai Yama
173. Margret Yama
174. Saraya Yanga
175. Yana Yidau
176. Hauwa Yirma
177. Naomi Zakaria

* = Spelling unconfirmed; names respectfully corrected to most likely spelling. Retyped by @ElisabethEpps.

15 April, 2014: At least 276 girls aged 15-18 yo were kidnapped from their boarding school in Chibok, Nigeria.
4 May, 2014: Christian Association of Nigeria released this list of 177 names reportedly confirmed still missing.

Systematic Sampling

A method of sampling items that involves selecting every k^{th} item in the population after randomly selecting a starting point between 1 and k. The value k is determined as the ratio of the population size over the desired sample size.

Advantages:

- Very easy to get sample

Disadvantages:

- May be biased, especially if order of list of population matters.

Example

A large worldwide financial company wants to develop a new retirement plan for the company. They want to survey different managers of branches around the world to find out the most important strategies the new retirement plan should contain. They have 5000 branches worldwide and want to personally interview these branch managers. They have information about the branch size (small, medium, large), and the state/province location of the branch. They want to talk to 50 branch managers.

Develop four separate strategies to sample these branch managers based on the four different statistical sampling techniques discussed previously.

Example

Develop four separate strategies to sample these branch managers based on the four different statistical sampling techniques discussed previously.

1. SRS – Randomly sample 50 branches to interview their managers
2. STS – Stratify by size and select SRS from each
3. Cluster – Randomly select sample of states/provinces, then interview branches from those states/provinces
4. Systematic – Select every 100^{th} branch in list of branches

Types of Data

4 Types of Data

There are four main types of data people typically deal with in data analysis.

These four types are split into two groups

1. Qualitative vs. Quantitative
2. Time Series vs. Cross-sectional

Quantitative vs. Qualitative

Quantitative:

- Data that are numeric that define value or quantity.
- Easy check → Must be able to do basic arithmetic and have it make sense.

Qualitative:

- Data whose measurement scale is inherently categorical.
- **Nominal** – categories with no logical ordering
- **Ordinal** – categories with a logical order / only two ways to order the categories (binary IS ordinal)

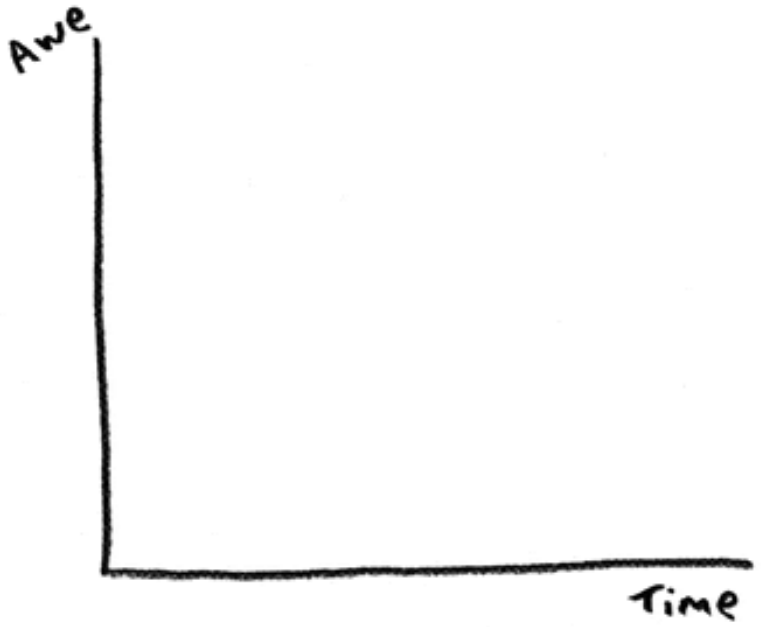
Time Series vs. Cross-sectional

Time Series:

- Set of ordered data values observed at successive points in time.

Cross-sectional:

- Set of data values observed at a fixed point in time, or where time is of no significance.



How I feel about mountains