[29]:	Data columns # Column 0 id 1 pH 2 Hardness 3 Solids 4 Chlorami 5 Sulfate 6 Conducti 7 OrganicC 8 Trihalom 9 Turbidit 10 Potabili dtypes: float memory usage:  display(Marko no1 = df.drop  stat = pd.Dat stat["Mean"] stat["Median" stat["Std"] = stat["Var"] = stat["Min"] = stat["Q1"] = stat["Q2"] = stat["Q2"] = stat["IQR"] = stat["Skew"] stat["Kurt"] stat	2010 non-null 20	nt Dtype int64 I float64 I float67 I float64 I flo	Minimum, Max Var				7.0294
[30]:	Chloramines Sulfate Conductivity OrganicCarbon Trihalomethanes Turbidity	195.969209 197.203525 21904.673439 20926.882155 7.134322 7.142014 333.211376 332.214113 426.476708 423.438372 14.357940 14.323286 66.400717 66.482041 3.969497 3.967374  down("#### Modus [*mul-	1.585214 2 41.211111 1 80.701872 6 3.325770 1 16.081109 2 0.780471 6	.439749e+07 320.9 .512904e+00 1.3 .698356e+03 129.0 .512792e+03 201.6 .106075e+01 2.2 .586021e+02 8.5 .091350e-01 1.4	90871 13.127000 00000 481.030642	11.736129 352.030642	176.740657 514.412962 2 6.138326 807.626986 866.619219 12.122530 55.949993 3.442882	197.2035 0926.8821 7.1420 332.2141 423.4383 14.3232 66.4820 3.9673
30]:	Modus [multive ph	Hardness         Solids           73.492234         320.942611           77.459586         1198.943699           81.710895         1351.906979           94.091307         1372.091043           94.812545         2552.962804               286.567991         50793.898917           287.975540         53735.899194           300.292476         55334.702799           306.627481         56351.396304           317.338124         56488.672413           olumns	1.920271 180 2.397985 182 2.456014 187 2.458609 187  12.580026 458 12.626900 460 12.653362 475 13.043806 476 13.127000 481	201.619 2.306746 210.319 2.397370 233.907 7.170714 245.859 7.424131 252.968 3.441072 666.690 2.107069 669.725 5.737460 695.369 5.539717 708.226	4.371899 4.466772 4.861631 4.966862 618 23.569645 086 23.604298 528 23.917601 364 24.755392 520 27.006707	114.034946 114.208671 116.161622 120.030077 124.000000	Turbidity  1.450000  1.492207  1.496101  1.680554  1.812529   6.307678  6.357439  6.389161  6.494249  6.494749	
	plt.title plt.show(	nt visualisasi plot te		14 - 12 - 10 - 8 - 6 - 4 - 2 -	uk histogram d	an boxplot	untuk s	etiap
[33]:	banyak titik outle no2 ("Hardness 600 - 500 - 200 -	togram yang terbentuk, terli arkan box plot yang terbent iers yang memiliki nilai lebih s")  Hardness	uk, terlihat bahw	va rentang antara	titik minimum dan ma	ksimum cukup l	cecil serta te	,
34]:	skewed). Berdasa	togram yang terbentuk, terli arkan box plot yang terbent iers yang memiliki nilai lebih Solids	uk, terlihat bahw	va rentang antara	titik minimum dan ma	ksimum cukup l	oesar serta te	
	Berdasarkan hist skewed). Berdasa	togram yang terbentuk, terli arkan box plot yang terbent iers yang memiliki nilai lebih ines")  Chloramines	50000 hat bahwa diagi uk, terlihat bahw	va rentang antara maksimum.		_	-	•
	arah kanan (no s terdapat banyak no2 ("Sulfate" 700 - 500 -	togram yang terbentuk, terliskew). Berdasarkan box plot a titik <i>outliers</i> yang memiliki	yang terbentuk,	terlihat bahwa re	ntang antara titik mini	mum dan maksi	mum cukup	kecil sert
37]:	Berdasarkan hist skewed). Berdasa	togram yang terbentuk, terli arkan box plot yang terbent <i>iers</i> yang memiliki nilai lebih	uk, terlihat bahv	va rentang antara ninimum atau me	titik minimum dan ma	ksimum cukup l	oesar serta te	•
	Berdasarkan hist skewed). Berdasa	togram yang terbentuk, terliarkan box plot yang terbentuk nilai lebat nilai nilai lebat nilai lebat nilai lebat nilai lebat nilai ni	uk, terlihat bahv	va rentang antara k maksimum.		_	•	•
	arah kanan (no s		yang terbentuk,	terlihat bahwa re	ntang antara titik mini m atau memiliki nilai l	mum dan maksi	mum cukup	kecil sert
	arah kanan (no s	togram yang terbentuk, terliskew). Berdasarkan box plot	yang terbentuk,	terlihat bahwa re	ntang antara titik mini	mum dan maksi	mum cukup	besar
[40]:	100 - 100 - 2	Turbidity 5	6	6 - 5 - 4 - 3 - 2 -	Turbidity			
[41]:	arah kanan (no s terdapat sejumla  3. Menentu  import scipy. import warning warnings.simp  def no3(kolom stat, p = display(M if (p > 0 displ else: displ plt.subpl df[kolom] df[kolom] plt.title plt.show(	ngs plefilter(action="igno:  m): = st.normaltest(df[kolo Markdown("D'Agostino's ).05): lay(Markdown("Kolom %s lay(Markdown("Kolom %s lots()  .plot.hist(density="T:  .plot.kde(color="tab:)e(kolom)	yang terbentuk, ki nilai lebih keci numerik be re", category  DM]) K-squared Te Berdistribus Tidak Berdis	terlihat bahwa re I dari titik minimu erdistribusi n =FutureWarning est, \$p\$ = %.10 i Normal" %(kc	ntang antara titik mini m atau memiliki nilai l ormal atau tida )  f" % (p)))  lom)))	mum dan maksi ebih besar dari t	mum cukup	kecil sert
[44]:	Nolom pH Tidak R  0.25  0.20  0.10  0.05  0.00  1.5  0.00  1.5  0.7  0.7  0.7  0.7  0.7  0.7  0.7  0	quared Test, $p$ = 0.00002651  Berdistribusi Normal  pH  quared Test, $p$ = 0.00013442  Tidak Berdistribusi Normal	15 20					
	-		300 400					
	-							
	0.10  0.05  0.00  -5  no3 ("Sulfate" D'Agostino's K-so	quared Test, $p = 0.00000044$ dak Berdistribusi Normal Sulfate	226	20				
	-	quared Test, $p = 0.00000043$ vity Tidak Berdistribusi Norm  Conductivity	90	00				
50]:	O.10 - O.08 - O.04 - O.02 - O.00 - O.	quared Test, $p$ = 0.10435984 thanes Berdistribusi Normal	30 4	10				
	3	Trihalomethanes  50  100  2y")  quared Test, $p = 0.76947173$ Berdistribusi Normal  Turbidity	150					
[52]:	import scipy. from statsmood from statsmood  a. Nilai rata-ra: $H_0$ : Nilai rata-ra: $H_1$ : Nilai rata-ra: Tingkat Signifika	dels.stats.weightstats dels.stats.proportion : $lpha$ at a pH di at as 7? ta pH sama dengan 7 ( $\mu = lpha$ ta pH lebih dari 7 ( $\mu > 7$ ) an: $lpha = 0.05$	import ztest import propor  7)	tions_ztest				
	Rumus perhitung Jika $z$ berada pa	da $critical$ $section$ $(z>z_lpha)$ , da $pada$ $critical$ $section$ $(z\le c (kind="box")$	maka nilai $p < lpha$	$z=rac{ar{x}-\mu_0}{\sigma/\sqrt{n}}$ $lpha$ dan $H_0$ ditolak	ak ditolak			
	display (Marko  z_alpha = st. display (Marko  if z > z_alph	o pH  (df["pH"], value=miu_0 down(f"Didapatkan nila norm.ppf(1 - alpha) down(f"Didapatkan nila na: Markdown(f"Keputusan: Markdown(f"Tes Signifi Markdown(f"Jadi, nilai Markdown(f"Tes Signifi Markdown(f"Tes Signifi Markdown(f"Tes Signifi Markdown(f"Tes Signifi Markdown(f"Tes Signifi Markdown(f"Jadi, nilai  z = 2.4854 dan p = 0.006	i \$z = {round i \$z_\\alpha  Folak \$H_0\$ d kan: Tolak \$H rata-rata pH  Fidak menolak kan: Tidak me rata-rata pH	<pre>(z, 4)}\$ dan \$ = {round(z_alp ikarenakan \$z\$</pre>	ha, 4)}\$"))  berada pada *cri n \$p\$ lebih kecil miu_0}\$"))  akan \$z\$ tidak be karenakan \$p\$ leb	tical sectior dari tingkat rada pada *cı	signifik	an yang
	Jadi, nilai rata-rat $H_0$ : Nilai rata-ra $H_1$	gle Sample Two Tailed Z Test ${ m gan\ nilai\ }z$ : ${ m da\ }critical\ section\ (z<-z_lpha$ ${ m da\ }p$ ${ m ada\ }critical\ section\ (z<$	ma dengan 20 $205~(\mu=205)$ gan $205~(\mu\neq2)$ dengan $205~(\mu\neq2)$ dengan $205~(\mu\neq2)$	$z=rac{ar{x}-\mu_0}{\sigma/\sqrt{n}}$ (2), maka nilai $p<0$	$x_2$ atau $z>z_{lpha/2}$ $lpha$ dan $H_0$ ditolak	ditolak		
	<pre> AxesSubplot:  300 -  250 -  200 -  150 -  100 -  100 -  miu_0 = 205 alpha = 0.05  z, p = ztest( display(Marko) display(Marko)  if z &lt; -z_alph</pre>	Hardness  (df["Hardness"], valued down(f"Didapatkan niladdua = st.norm.ppf(1 - addown(f"Didapatkan niladdua = st.norm.ppf(1 - addown(f"Tes Signific Markdown(f"Tes Signific Markdown(f"Tes Signific Markdown(f"Tes Signific Markdown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Tes Signific Markdown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Tes Signific Markdown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Tes Signific Markdown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Tes Signific Markdown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Tes Signific Markdown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Tes Signific Markdown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Jadi, niladdua = st.norm.ppf(1 - addown(f"Jad	<pre>i \$z = {round alpha / 2) i \$z_{{\\alpha} lpha_per_dua: Folak \$H_0\$ d kan: Tolak \$H   rata-rata Ha Fidak menolak kan: Tidak me   rata-rata Ha</pre>	<pre>(z, 4)}\$ dan \$ a/2}} = {round ikarenakan \$z\$ _0\$ dikarenaka rdness tidak s  \$H_0\$ dikaren nolak \$H_0\$ di</pre>	<pre>p = {round(p, 4)} (z_alpha_per_dua,  berada pada *cri n \$p\$ lebih kecil ama dengan \${miu_ akan \$z\$ tidak be karenakan \$p\$ leb</pre>	4)}\$"))  tical section   dari tingkat 0}\$"))  rada pada *cr	signifik	an yang
	Didapatkan nilai $z$ Keputusan: Tolak Tes Signifikan: Tol Jadi, nilai rata-rat <b>c. Nilai rata-r</b> $H_0$ : Nilai rata-rat  Tingkat Signifika  Uji Statistik: Sing  Rumus perhitung  Jika $z$ berada pa  Jika $z$ tidak berad $df$ ["Solids"]	$z_{lpha/2}=1.96$ $H_0$ dikarenakan $z$ berada palak $H_0$ dikarenakan $p$ lebih ta Hardness tidak sama dengata 100 baris pertama kata 100 baris pertama kolomata 100 baris pertama k	kecil dari tingkatgan $205$ $ extbf{kolom Solids}$ Solids sama der Solids tidak sama dengan $Critical$ $ extbf{dengan Critical}$ $ extbf{dengan Critical}$ $ extbf{dengan Critical}$ $ extbf{dengan Critical}$	t signifikan yang o $z$ bukan 21900? $z$ ngan 21900 ( $\mu=1$ na dengan 21900 $z=1$ $z=$	iinginkan ( $p < 0.05$ ) $21900$ ) $(\mu  eq 21900)$ $a_2$ atau $z > z_{lpha/2}$ $lpha$ dan $H_0$ ditolak	ditolak		
[57]:   [57]:	<axessubplot:< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td></axessubplot:<>							

	700 - 600 - 500 - 400 - 300 - 200 -	Conductivity
60]:	alph z, p disp z_al disp if z	l = 0.10 a = 0.05  = proportions_ztest(len(df[df["Conductivity"] > 450]), len(df), value=p_nol, alternative="two-sided", lay(Markdown(f"Didapatkan nilai \$z = {round(z, 4)}\$ dan \$p = {round(p, 4)}\$"))  pha_per_dua = st.norm.ppf(1 - alpha / 2) lay(Markdown(f"Didapatkan nilai \$z_{{\alpha/2}} = {round(z_alpha_per_dua, 4)}\$"))  <-z_alpha_per_dua or z > z_alpha_per_dua: display(Markdown(f"Keputusan: Tolak \$H_0\$ dikarenakan \$z\$ berada pada *critical section* (\$z < {round(display(Markdown(f"Tes Signifikan: Tolak \$H_0\$ dikarenakan \$p\$ lebih kecil dari tingkat signifikan yang display(Markdown(f"Jadi, persentase nilai Conductivity yang lebih dari \$450\$ adalah tidak sama dengan \$c:
  -  -	Didap Didap Keput Tes Sig Jadi, p	display (Markdown (f"Keputusan: Tidak menolak $H_0$ \$ dikarenakan $z$ \$ tidak berada pada *critical section* display (Markdown (f"Tes Signifikan: Tidak menolak $H_0$ \$ dikarenakan $z$ \$ lebih besar atau sama dengan tirdisplay (Markdown (f"Jadi, persentase nilai Conductivity yang lebih dari $z$ 450\$ adalah sama dengan $z$ 0 atkan nilai $z$ 40.4464 dan $z$ 40 dikarenakan $z$ 40 dikarenakan $z$ 40 berada pada critical section ( $z$ 40 dikarenakan $z$ 50 dia dalah tidak sama dengan 0.1 dikarenakan $z$ 50 dikarenakan $z$ 50 dikarenakan dengan 0.1 dikarenakan $z$ 50 dikarenakan $z$ 50 dikarenakan dengan 0.1 dikarenakan $z$ 50 dikarenakan dikarenakan dengan 0.1 dikarenakan dikarenakan dikarenakan dikarenakan dikarenakan dikarenakan dengan 0.1 dikarenakan dikarenakan dikarenakan dikarenakan dikarenakan dikarenakan dengan 0.1 dikarenakan dikarenakan dikarenakan dengan 0.1 dikarenakan
	$H_1$ : For Tingk  Uji St  Rumu	Persentase nilai Conductivity yang lebih dari $450$ adalah sama dengan $5\%$ $(p_0=0.05)$ Persentase nilai Conductivity yang lebih dari $450$ adalah kurang dari $5\%$ $(p_0<0.05)$ Rat Signifikan: $\alpha=0.05$ Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $Critical$ $Section$ $z< z_{\alpha}$ Rus perhitungan nilai $z$ : $z=\frac{\hat{p}-p_0}{\sqrt{p_0q_0/n}}$ Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 0.05 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 1.05 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 2.06 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 3.07 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 3.07 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 3.07 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 3.07 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 4.07 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 4.07 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 4.08 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 4.09 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 5.09 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 6.09 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 6.09 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 6.09 Ratistik: Single Sample Left Tailed One Proportion Z Test dengan $z$ 7.09 Ratistik: Single Sample
	Jika z	tidak berada pada $critical$ $section$ ( $z \geq z_{lpha}$ ), maka nilai $p \geq lpha$ dan $H_0$ tidak ditolak
2]:	z, p disp z_al	Tihalomethanes  1 = 0.05 a = 0.05  = proportions_ztest(len(df[df["Trihalomethanes"] < 40]), len(df), value=p_nol, alternative="smaller", lay(Markdown(f"Didapatkan nilai \$z = {round(z, 4)}\$ dan \$p = {round(p, 4)}\$"))  pha = st.norm.ppf(1 - alpha) lay(Markdown(f"Didapatkan nilai \$z_{{\alpha}} = {round(z_alpha, 4)}\$"))
1	else Didap Didap Keput	<pre>&lt;-z_alpha: display(Markdown(f"Keputusan: Tolak \$H_0\$ dikarenakan \$z\$ berada pada *critical section* (\$z &lt; {round(-display(Markdown(f"Tes Signifikan: Tolak \$H_0\$ dikarenakan \$p\$ lebih kecil dari tingkat signifikan yang display(Markdown(f"Jadi, persentase nilai Trihalomethanes yang kurang dari \$40\$ adalah kurang dari \${p_0}: display(Markdown(f"Keputusan: Tidak menolak \$H_0\$ dikarenakan \$z\$ tidak berada pada *critical section* display(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tir display(Markdown(f"Jadi, persentase nilai Trihalomethanes yang kurang dari \$40\$ adalah sama dengan \${p_0} atkan nilai <math>z = 0.5629 \text{ dan } p = 0.7132</math> atkan nilai <math>z = 0.5629 \text{ dan } p = 0.7132</math> grifikan: Tidak menolak <math>H_0</math> dikarenakan <math>p</math> lebih besar atau sama dengan tingkat signifikan yang diinginkan (<math>p \ge 0.05</math>)</pre>
	Jadi, p  5. N  a. Daterse $H_0$ : N $H_1$ : N	Persentase nilai Trihalomethanes yang kurang dari $40$ adalah sama dengan $0.05$ Melakukan test hipotesis $2$ sampel.  Ata kolom Sulfate dibagi sama rata: bagian awal dan bagian akhir kolom. Benarkah rata-rata kedua bagian ebut sama?  Nilai rata-rata kolom awal Sulfate sama dengan nilai rata-rata kolom akhir Sulfate ( $\mu_1 - \mu_2 = 0$ )  Nilai rata-rata kolom awal Sulfate tidak sama dengan nilai rata-rata kolom akhir Sulfate ( $\mu_1 - \mu_2 \neq 0$ )  Kat Signifikan: $\alpha = 0.05$
	Rumu Jika <i>z</i> Jika <i>z</i>	atistik: Double Sample Two Tailed Z Test dengan $Critical$ $Section$ $z<-z_{\alpha/2}$ atau $z>z_{\alpha/2}$ atau $z>z_{\alpha/2}$ at $z>z_{\alpha/2}$ , maka nilai $z>z$ at au $z>z_{\alpha/2}$ at au $z>z_{\alpha/2}$ , maka nilai $z>z$ at au $z>z_{\alpha/2}$ .
3]:	450 - 400 - 350 - 300 - 250 - 200 -	sSubplot:>
64]:		Sulfate"][len(df) // 2:].plot(kind="box") sSubplot:>
55]:	<pre>z, p disp z_al disp</pre>	Sulfate  a_miu = 0 a = 0.05  = ztest(df["Sulfate"][:len(df) // 2], df["Sulfate"][len(df) // 2:], value=delta_miu, alternative="two-lay(Markdown(f"Didapatkan nilai \$z = {round(z, 4)}\$ dan \$p = {round(p, 4)}\$"))  pha_per_dua = st.norm.ppf(1 - alpha / 2)
  -  -	else Didap Didap Keput Tes Sig	display (Markdown (f"Keputusan: Tolak $$H_0$$$ dikarenakan $$z$$ berada pada *critical section* ( $$z < {round (display (Markdown (f"Tes Signifikan: Tolak $H_0$$ dikarenakan $p$ lebih kecil dari tingkat signifikan yang display (Markdown (f"Jadi, nilai rata-rata kolom awal Sulfate tidak sama dengan nilai rata-rata kolom akhi: display (Markdown (f"Keputusan: Tidak menolak $H_0$$ dikarenakan $z$ tidak berada pada *critical section* display (Markdown (f"Tes Signifikan: Tidak menolak $H_0$$ dikarenakan $p$ lebih besar atau sama dengan tir display (Markdown (f"Jadi, nilai rata-rata kolom awal Sulfate sama dengan nilai rata-rata kolom akhir Sulatkan nilai z=-2.0753 dan p=0.038 atkan nilai z=-2.0753 dan z=0.038 dikarenakan z=$
	b. Do awal $H_0$ : F $\mu_1$ — $H_1$ : F $\mu_1$ — Tingk	ata kolom OrganicCarbon dibagi sama rata: bagian awal dan bagian akhir kolom. Benarkah rata-rata bagia l lebih besar daripada bagian akhir sebesar 0.15? Perbedaan nilai rata-rata kolom awal OrganicCarbon dengan nilai rata-rata kolom akhir OrganicCarbon sama dengan 0.15 ( $\mu_2=0.15$ ) Perbedaan nilai rata-rata kolom awal OrganicCarbon dengan nilai rata-rata kolom akhir OrganicCarbon tidak sama dengan 0.15 ( $\mu_2\neq0.15$ ) Perbedaan nilai rata-rata kolom awal OrganicCarbon dengan nilai rata-rata kolom akhir OrganicCarbon tidak sama dengan 0.15 ( $\mu_2\neq0.15$ ) Rat Signifikan: $\alpha=0.05$ Ratistik: Double Sample Two Tailed Z Test dengan $Critical$ $Section$ $z<-z_{\alpha/2}$ atau $z>z_{\alpha/2}$
56]:	Jika z Jika z df ["	$z=\frac{(\bar{x_1}-\bar{x_2})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$ z berada pada $critical$ $section$ $(z<-z_{\alpha/2}$ atau $z>z_{\alpha/2}$ ), maka nilai $p<\alpha$ dan $H_0$ ditolak z tidak berada pada $critical$ $section$ $(z<-z_{\alpha/2}$ atau $z>z_{\alpha/2}$ ), maka nilai $p\geq\alpha$ dan $H_0$ tidak ditolak OrganicCarbon"] [:len(df) // 2].plot(kind="box")
66]:	25	sSubplot:>
57]:   57]:		OrganicCarbon"][len(df) // 2:].plot(kind="box") sSubplot:>
58]:	<pre>z, p disp z_al disp</pre>	OrganicCarbon  a_miu = 0.15 a = 0.05  = ztest(df["OrganicCarbon"][:len(df) // 2], df["OrganicCarbon"][len(df) // 2:], value=delta_miu, alter clay(Markdown(f"Didapatkan nilai \$z = {round(z, 4)}\$ dan \$p = {round(p, 4)}\$"))  pha_per_dua = st.norm.ppf(1 - alpha / 2) clay(Markdown(f"Didapatkan nilai \$z_{{\alpha/2}} = {round(z_alpha_per_dua, 4)}\$"))
1	<b>else</b> Didap Didap Keput	<pre>&lt; -z_alpha_per_dua or z &gt; z_alpha_per_dua: display(Markdown(f"Keputusan: Tolak \$H_0\$ dikarenakan \$z\$ berada pada *critical section* (\$z &lt; {round(display(Markdown(f"Tes Signifikan: Tolak \$H_0\$ dikarenakan \$p\$ lebih kecil dari tingkat signifikan yang display(Markdown(f"Jadi, perbedaan nilai rata-rata kolom awal OrganicCarbon dengan nilai rata-rata kolom: display(Markdown(f"Keputusan: Tidak menolak \$H_0\$ dikarenakan \$z\$ tidak berada pada *critical section* display(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tir display(Markdown(f"Jadi, perbedaan nilai rata-rata kolom awal OrganicCarbon dengan nilai rata-rata kolom atkan nilai <math>z = -2.4131  \text{dan}  p = 0.0158</math> atkan nilai <math>z = -2.4131  \text{dan}  p = 0.0158</math> guifikan: Tolak <math>H_0</math> dikarenakan <math>z</math> berada pada <math>critical</math> <math>section</math> (<math>z &lt; -1.96</math> atau <math>z &gt; 1.96</math>) guifikan: Tolak <math>H_0</math> dikarenakan <math>p</math> lebih kecil dari tingkat signifikan yang diinginkan (<math>p &lt; 0.05</math>)</pre>
,	Jadi, p $\mu_1 - R$ <b>c.</b> $Ra$ $H_0$ : $N$ Tingk	perbedaan nilai rata-rata kolom awal OrganicCarbon dengan nilai rata-rata kolom akhir OrganicCarbon tidak sama dengan 0.15 ( $\mu_2 \neq 0.15$ )  ata-rata 100 baris pertama kolom Chloramines sama dengan 100 baris terakhirnya?  Nilai rata-rata 100 baris pertama kolom Chloramines sama dengan nilai rata-rata 100 baris kolom akhir Chloroamines ( $\mu_1 - \mu_2 = 0.000$ )  Nilai rata-rata 100 baris pertama kolom Chloramines sama dengan nilai rata-rata 100 baris kolom akhir Chloroamines ( $\mu_1 - \mu_2 \neq 0.000$ )  Sata Signifikan: $\alpha = 0.000$ Satistik: Double Sample Two Tailed Z Test dengan <i>Critical Section</i> $z < -z_{\alpha/2}$ atau $z > z_{\alpha/2}$
	Jika z Jika z df["	$z = \frac{(\bar{x_1} - \bar{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$ z berada pada $critical$ $section$ $(z < -z_{\alpha/2}$ atau $z > z_{\alpha/2}$ ), maka nilai $p < \alpha$ dan $H_0$ ditolak z tidak berada pada $critical$ $section$ $(z < -z_{\alpha/2}$ atau $z > z_{\alpha/2}$ ), maka nilai $p \ge \alpha$ dan $H_0$ tidak ditolak Chloramines"] .head (100) .plot (kind="box") s Subplot:>
	10 - 9 - 8 - 7 - 6 - 5 - 4 -	Chloramines
70]: 70]:	Axe  10 - 9 - 8 - 7 - 6 -	Chloramines"].tail(100).plot(kind="box") sSubplot:>
71]:	<pre>alph z, p disp z_al</pre>	Chloramines  a_miu = 0 a = 0.05  = ztest(df["Chloramines"].head(100), df["Chloramines"].tail(100), value=delta_miu, alternative="two-s: clay(Markdown(f"Didapatkan nilai \$z = {round(z, 4)}\$ dan \$p = {round(p, 4)}\$"))  pha_per_dua = st.norm.ppf(1 - alpha / 2) clay(Markdown(f"Didapatkan nilai \$z_{{\alpha/2}} = {round(z_alpha_per_dua, 4)}\$"))
	<b>else</b> Didap Didap Keput	display (Markdown (f"Keputusan: Tidak menolak \$H_0\$ dikarenakan \$z\$ tidak berada pada *critical section* display (Markdown (f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tindisplay (Markdown (f"Jadi, Nilai rata-rata 100 baris pertama kolom Chloramines sama dengan nilai rata-rata atkan nilai $z=-0.7059$ dan $p=0.4802$ atkan nilai $z_{\alpha/2}=1.96$ usan: Tidak menolak $H_0$ dikarenakan $z$ tidak berada pada $critical$ $section$ $(-1.96 \le z \le 1.96)$
,	Jadi, N $H_0$ : F $h_1$ : F $h_2$ : F $h_3$ : F $h_4$ :	gnifikan: Tidak menolak $H_0$ dikarenakan $p$ lebih besar atau sama dengan tingkat signifikan yang diinginkan ( $p \geq 0.05$ ) Nilai rata-rata 100 baris pertama kolom Chloramines sama dengan nilai rata-rata 100 baris kolom akhir Chloroamines ( $\mu_1 - \mu_2 = 0$ ) Proporsi nilai bagian awal Turbidity yang lebih dari 4 adalah lebih besar daripada proporsi nilai yang sama adi bagian akhir Turbidity? Proporsi nilai bagian awal Turbidity yang lebih dari 4 sama besar dengan proporsi nilai yang sama pada di bagian akhir Turbidity ( $p_2 = 0$ ) Proporsi nilai bagian awal Turbidity yang lebih dari 4 lebih besar dengan proporsi nilai yang sama pada di bagian akhir Turbidity ( $p_2 = 0$ )
	Uji St Rumu Jika 2	tat Signifikan: $\alpha=0.05$ atistik: Single Sample Right Tailed One Proportion Z Test dengan $Critical$ $Section ~z>z_{lpha}$ as perhitungan nilai $z$ : $z=\frac{\hat{p_1}-\hat{p_2}}{\sqrt{\hat{p}\hat{q}\left(1/n_1+1/n_2\right)}}$ as berada pada $critical$ $section~(z>z_{lpha})$ , maka nilai $p<\alpha$ dan $H_0$ ditolak atidak berada pada $critical$ $section~(z\leq z_{lpha})$ , maka nilai $p\geq \alpha$ dan $H_0$ tidak ditolak
/2]: [ /2]:		Turbidity"][:len(df) // 2].plot(kind="box") sSubplot:>
/3]:   /3]:		Turbidity"][len(df) // 2:].plot(kind="box") sSubplot:>
74]:	delt alph	Turbidity  a_p = 0 a = 0.05  val = df[:len(df) // 2]
	<pre>dfAk z, p disp z_al disp if z</pre>	thir = df[len(df) // 2:]  = proportions_ztest([len(dfAwal[dfAwal["Turbidity"] > 4]), len(dfAkhir[dfAkhir["Turbidity"] > 4])], [lay(Markdown(f"Didapatkan nilai \$z = {round(z, 4)}\$ dan \$p = {round(p, 4)}\$"))  pha = st.norm.ppf(1 - alpha) lay(Markdown(f"Didapatkan nilai \$z_{{\alpha}} = {round(z_alpha, 4)}\$"))  > z_alpha: display(Markdown(f"Keputusan: Tolak \$H_0\$ dikarenakan \$z\$ berada pada *critical section* (\$z > {round(z_display(Markdown(f"Tes Signifikan: Tolak \$H_0\$ dikarenakan \$p\$ lebih kecil dari tingkat signifikan yang display(Markdown(f"Jadi, proporsi nilai bagian awal Turbidity yang lebih dari 4 lebih besar dengan propositical display(Markdown(f"Keputusan: Tidak menolak \$H_0\$ dikarenakan \$z\$ tidak berada pada *critical section* display(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$ dikarenakan \$p\$ lebih besar atau sama dengan tinglay(Markdown(f"Tes Signifikan: Tidak menolak \$H_0\$) dikarenakan \$p\$ lebih kecil dikarenakan \$p\$ leb
] - - 2	Didap Didap Keput Tes Sig Jadi, p $p_1-p_1$	display (Markdown (f"Jadi, proporsi nilai bagian awal Turbidity yang lebih dari 4 sama besar dengan proposatkan nilai $z=-0.1339$ dan $p=0.5533$ atkan nilai $z_{\alpha}=1.6449$ usan: Tidak menolak $H_0$ dikarenakan $z$ tidak berada pada $critical$ $section$ ( $z\leq 1.6449$ ) gnifikan: Tidak menolak $H_0$ dikarenakan $p$ lebih besar atau sama dengan tingkat signifikan yang diinginkan ( $p\geq 0.05$ ) proporsi nilai bagian awal Turbidity yang lebih dari 4 sama besar dengan proporsi nilai yang sama pada di bagian akhir Turbidity ( $p_2=0$ ) agian awal kolom Sulfate memiliki variansi yang sama pada bagian akhirnya?
	$H_1$ : N ${f Tingk}$ Uji St ${f v}_2=$	/ariansi pada bagian awal kolom Sulfate sama dengan variansi pada bagian akhir kolom Sulfate $({\sigma_1}^2={\sigma_2}^2)$ /ariansi pada bagian awal kolom Sulfate tidak sama dengan variansi pada bagian akhir kolom Sulfate $({\sigma_1}^2\neq{\sigma_2}^2)$ sat Signifikan: $\alpha=0.05$ satistik: Double Sample Two Tailed F Test dengan <i>Critical Section</i> $f< f_{1-\alpha/2}(v_1,v_2)$ atau $f>f_{\alpha/2}(v_1,v_2)$ dengan $v_1=n_1-1$ da $n_2-1$ sus perhitungan nilai $f$ : $f=\frac{s_1^2}{s_2^2}$
	Jika j	f berada pada $critical$ $section$ $(f < f_{1-\alpha/2}(v_1,v_2)$ atau $f > f_{\alpha/2}(v_1,v_2)$ ), maka nilai $p < \alpha$ dan $H_0$ ditolak f tidak berada pada $critical$ $section$ $(f < f_{1-\alpha/2}(v_1,v_2)$ atau $f > f_{\alpha/2}(v_1,v_2)$ ), maka nilai $p \ge \alpha$ dan $H_0$ tidak ditolak $Sulfate"$ ][:len(df) // 2].plot(kind="box") $Subplot:>$
′6]: [ ′6]:		Sulfate  Sulfate"][len(df) // 2:].plot(kind="box") sSubplot:>
	450 - 400 - 350 - 300 - 250 - 200 -	Sulfate
	<pre>v_1 v_2 f = p = disp f_lo f_up disp disp</pre>	<pre>a = 0.05 = len(df[:len(df) // 2]) - 1 = len(df[len(df) // 2:]) - 1  df["Sulfate"][:len(df) // 2].var() / df["Sulfate"][len(df) // 2:].var() 1 - st.f.cdf(f, v_1, v_2)</pre>
 	else Didap Didap Didap Keput	display(Markdown(f"Jadi, variansi pada bagian awal kolom Sulfate tidak sama dengan variansi pada bagian
	6. N plot Deng 1. Ha 1. ( 2. ( 3. (	Aenentukan korelasi kolom non-target terhadap target dan membuat visualisasi scatter tnya.  Jan menggunakan tes korelasi, kita akan mendapatkan nilai Pearson Correlation Coefficient $\rho_{X,Y}$ yang rentangnya antara -1 sampa sil tes korelasi akan diklasifikasi menggunakan aturan praktis oleh Rea dan Parker (1992), yaitu: $0.00 < 0.10$ - Negligible $0.10 < 0.20$ - Weak $0.20 < 0.40$ - Moderate $0.40 < 0.60$ - Relatively strong
112	5. ( 6. ( <b>def</b>	0.60 < 0.80 - Strong 0.80 < 1.00 - Very strong  no6(kolom): rho, p_value = st.pearsonr(df[kolom], df["Potability"]) print("Korelasi", kolom, "dan target (Potability)") print("rho = ",rho, "\np-value = ", p_value) if (p_value < 0.05):     print("Kolom", kolom, "dan target berkorelasi")     if (rho >= 0.8):         print("Kolom", kolom, "dan target memiliki Very Strong Correlation") elif (rho >= 0.6):         print("Kolom", kolom, "dan target memiliki Strong Correlation") elif (rho >= 0.4):
	no6(	<pre>print("Kolom", kolom, "dan target memiliki Relatively Strong Correlation") elif (rho &gt;= 0.2):     print("Kolom", kolom, "dan target memiliki Moderate Correlation") elif (rho &gt;= 0.1):     print("Kolom", kolom, "dan target memiliki Weak Correlation") else:     print("Kolom", kolom, "dan target memiliki Negligible Correlation") else:     print("Kolom", kolom, "dan target tidak berkorelasi") plt.scatter(df[kolom], df["Potability"]) plt.show()  "pH") lasi pH dan target (Potability)</pre>
	rho p-va	= 0.015475094408433502 lue = 0.48805529843977147 m pH dan target tidak berkorelasi
	no6( Kore rho p-va	"Hardness")  lasi Hardness dan target (Potability) = -0.0014631528959479423 lue = 0.9477306859719101 m Hardness dan target tidak berkorelasi
	Kore	"Solids") lasi Solids dan target (Potability) = 0.03897657818173474
	p-va	lue = 0.08063479734364612 m Solids dan target tidak berkorelasi
	Kore rho p-va Kolo	0 10000 20000 30000 40000 50000  "Chloramines")  lasi Chloramines dan target (Potability)  = 0.02077892184052413  lue = 0.3518003359635207  m Chloramines dan target tidak berkorelasi
	Kore	"Sulfate") lasi Sulfate dan target (Potability) = -0.015703164419273812 lue = 0.48166585681668234
	0.8 - 0.6 - 0.4 - 0.2 -	lue = 0.48166585681668234 m Sulfate dan target tidak berkorelasi
	Kore rho p-va Kolo	"Conductivity")  lasi Conductivity dan target (Potability) = -0.016257120111377105 lue = 0.4663378059686175 m Conductivity dan target tidak berkorelasi
	no6(	200 300 400 500 600 700  "OrganicCarbon")  lasi OrganicCarbon dan target (Potability)  = -0.015488461910747308
	p-va	= -0.015488461910747308 lue = 0.4876795482205311 m OrganicCarbon dan target tidak berkorelasi
	no6( Kore rho p-va	"Trihalomethanes")  lasi Trihalomethanes dan target (Potability)  = 0.009236711064713042  lue = 0.6789759745975009  m Trihalomethanes dan target tidak berkorelasi
	0.4 - 0.2 - 0.0 - no6 (	Z0 40 60 80 100 120  "Turbidity")  lasi Turbidity dan target (Potability) = 0.0223310426406227 lue = 0.316986307538362
	rho p-va	= 0.0223310426406227
	0.6 -	