

Laporan Projek 4

Literasi Liberal Arts

Kelompok 8:

- Bryan Christopher Wijaya
- James Patrick Oentoro
- Noel Christevent Mandak
- Stefannus Christian
- Tiffany Sondakh

Link tugas pemrograman :

<https://datalore.jetbrains.com/notebook/CPsQxEPW7yVpS665a3lC8b/HQnffFKXGwSjP7woAxXKRS/>

Masalah 1

Benarkah bahwa secara keseluruhan persentase mahasiswa CIT yang tidak meleak Liberal Arts lebih besar dari 75%?

Masalah 2

Benarkah bahwa secara keseluruhan, besarnya persentase mahasiswa CIT yang tidak pernah sama sekali di sepanjang semester untuk meminjam atau membaca buku non-prodinya di perpustakaan lebih besar ketika dievaluasi pada semester ganjil ketimbang pada semester genap?

Masalah 3

Apakah benar bahwa mahasiswa CIT mengalami peningkatan literasi Liberal Arts ketika membandingkan pada pengalamannya di tahun pertama dan setelah tahun pertama?

1. Sebagai pengantar, tulislah ringkasan singkat (maksimal 4 paragraf) untuk menceritakan rumusan masalah dan proses statistika yang Anda lakukan untuk menyelesaikan masalah.
2. Pada bagian inti, tuliskan secara mendetail mengenai bagaimana hasil uji hipotesis Anda dapat menyelesaikan setiap masalah yang diberikan.
3. Pada bagian penutup, tuliskan mengenai kesimpulan dan saran Anda dari hasil pengerjaan proyek.

- Ringkasan Singkat

- a. Rumusan Masalah:

- i. Dr. Calvin memiliki kekhawatiran mengenai literasi Liberal Arts pada mahasiswa CIT. Literasi Liberal Arts yang dimaksudkan adalah saat mahasiswa memiliki rasa ingin tahu terhadap hal-hal di luar bahasa program studinya sendiri. Dr. Calvin merasa sangat prihatin jika ternyata sebagian besar mahasiswa CIT tidak pernah datang ke perpustakaan untuk membaca ataupun meminjam buku

yang di luar cakupan Prodinya. Dari kekhawatiran Dr. Calvin, terdapat 3 rumusan masalah. Dr. Calvin meminta kami untuk menguji dugaan-dugaan tersebut secara statistika. 3 Rumusan masalah tersebut adalah:

- 1) Apakah keseluruhan persentase mahasiswa CIT yang **tidak** melek RLAC $\geq 75\%$.
- 2) Apakah besarnya persentase mahasiswa CIT yang tidak pernah sama sekali di sepanjang semester untuk meminjam atau membaca buku non-prodinya di perpustakaan lebih besar di semester ganjil ketimbang pada semester genap.
- 3) Apakah Mahasiswa CIT mengalami peningkatan literasi RLAC ketika membandingkan pada pengalamannya di tahun pertama dan setelah tahun pertama.

b. Proses Statistika

- i. Hal pertama yang dilakukan adalah mengumpulkan data sampel. Kami meminta data sampel dari pihak perpustakaan CIT secara langsung. Hal ini menyebabkan data sampel yang kami miliki untuk membuat hipotesis adalah data sampel yang saling bebas dan otentik.
- ii. Kemudian kami melakukan data preprocessing dengan cara menambahkan kolom Buta RLAC. Kolom ini akan memberikan nilai 1 jika orang tersebut membaca buku yang non - RLAC atau membaca buku sesuai dengan prodinya, dan akan memberikan nilai 0 jika orang tersebut membaca buku RLAC atau membaca buku di luar prodinya. Langkah ini membuat suatu masalah yaitu jika misalnya orang tersebut meminjam atau membaca suatu buku lebih dari sekali. Katakanlah misal orang A pernah membaca buku non prodi dan prodinya. Maka data orang ini akan menjadi bias. Untuk mengatasi hal ini, kami menggunakan fungsi group by berdasarkan nama dan mengambil data yang minimal. Setelah dilakukan proses ini, maka data yang tersisa adalah 0 saja (Melek RLAC) karena kami mengambil yang minimal setelah di group by berdasarkan nama.

Kode Pra-Pemrosesan Data

```
1 df1 = df.groupby("Nama Anggota").min()
2 df1
```

	ID Anggota	Kode Eksemplar	Judul	Tanggal Pinjam	Tanggal Kembali	Status peminjaman	Buta RLAC	angkatan
Nama Anggota								
Aaron Arquette	212100246	1033	Kalkulus, Jilid 1 = Calculus	10/1/2021	10/1/2021	1	0	21
Abrielle Irene Mirah	212100256	2147	Morfologi Tumbuhan	1/26/2022	2/9/2022	1	1	21
Adelaide Astrid Kusumawardhani	212100124	818	Brock Biology of Microorganisms	1/10/2022	1/24/2022	0	1	21
Alfa Marzelino	191900672	529	Depresi: Memandang Melampaui Kegelapan yang Su...	1/20/2022	10/12/2021	1	0	19
Alfred William Silele	212100709	2146	Morfologi Tumbuhan	2/2/2022	2/16/2022	1	1	21
...
Yabes Elmon	202000638	142	Allah Tritunggal	8/30/2021	9/13/2021	1	0	20
Yehezkiel Purnama Wijaya	10101190476	32	God's Fiery Challenger for Our Time: Festchrif...	8/25/2021	8/19/2021	1	0	19
Yolanda Charlota Thenu	202001260	746	Allah Tritunggal	1/25/2022	2/8/2022	0	0	20
Yosephine Priscilla Maris	191900292	735	Biology: A Global Approach	11/8/2021	1/12/2022	1	0	19
Yovan Carvallo	202000243	2500	Analisis Struktur dengan Metode Matriks Menggu...	10/22/2021	11/5/2021	1	1	20

162 rows × 8 columns

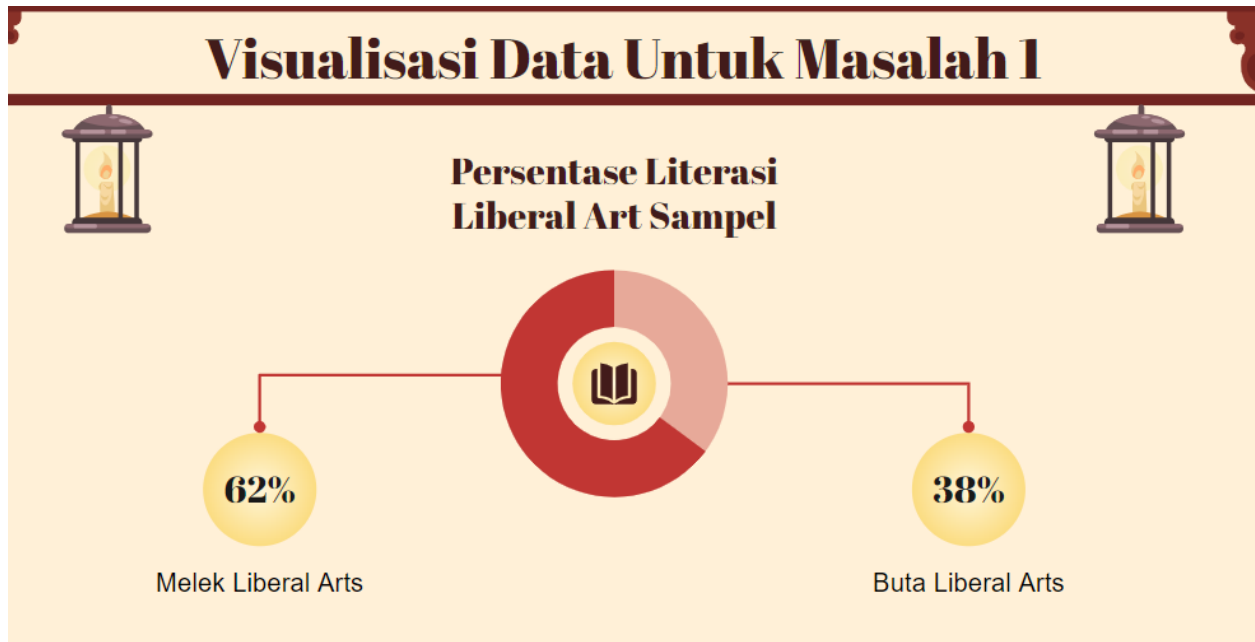
- iii. Setelah itu, kami menentukan hipotesis null dan hipotesis alternatif untuk ketiga masalah dari Dr. Calvin. Kami memiliki tingkat signifikansi 0.01 untuk ketiga masalah. Setelah itu kami memilih statistik yang cocok untuk membangun critical region atau p-value berdasarkan nilai alpha. Kemudian, kami membuat kesimpulan hasil hipotesis, kesimpulan detail akan dijelaskan di nomor 2. Langkah terakhir yang dilakukan adalah menerjemahkan kesimpulan ke dalam bahasa sehari-hari sehingga orang awam yang tidak mengerti statistika dapat mengerti kesimpulan yang dibuat.

- Hasil Uji Hipotesis

- **Masalah 1 :**

Benarkah bahwa secara keseluruhan persentase mahasiswa CIT yang tidak melek Liberal Arts lebih besar dari 75% ?

a. Visualisasi Data Masalah 1



Dapat dilihat dari visualisasi data diatas bahwa secara kasar, 62% dari data sampel Melek Liberal Arts dan 38% Buta Liberal Arts. Dapat ditebak bahwa dari data diatas bahwa seharusnya hasil statistika hipotesis yang didapatkan adalah bahwa mahasiswa Mahasiswa CIT yang tidak melek Liberal Arts tidak lebih dari 75%. Hal ini akan dibuktikan dengan langkah-langkah statistika dibawah ini.

Langkah pertama adalah menentukan hipotesis null dan hipotesis alternatif. Untuk masalah satu, kami memilih tingkat signifikansi 0.01 (tingkat kepercayaan 99%) dengan hipotesis null yaitu $H_0: \mu \leq 75\%$ dan $H_1: \mu > 75\%$.

Statistik yang dipakai dalam masalah satu adalah ratahan satu populasi.

Perhitungan Statistika. Kami melakukan perhitungan statistika dengan membagi ke dalam dua cara yaitu cara manual dan cara modern.

Perhitungan Cara Manual Masalah 1

```
1 x_bar = np.mean(df1['Buta RLAC'])
2 print("Rata-rata sampel yang buta RLAC:", x_bar)
3
4 n = len(df1)
5 print("Jumlah data sampel:", n)
6
7 std_dev = df1['Buta RLAC'].std()
8 print("Standar deviasi sampel yang buta RLAC:", std_dev)
9
10 alpha = 0.01
11 print("Tingkat kepercayaan:", 1-alpha)
```

```
Rata-rata sampel yang buta RLAC: 0.3765432098765432
Jumlah data sampel: 162
Standar deviasi sampel yang buta RLAC: 0.48602113926325335
Tingkat kepercayaan: 0.99
```

Karena sampel yang kami kumpulkan lebih besar dari 30, maka data sampel kami termasuk data sampel yang besar. Karena itu kami menggunakan distribusi normal (Z). Referensi rumus yang dipakai adalah sebagai berikut.

Statistik yang diuji	H_0	Statistik Uji	Distribusi	Kondisi	Method Python	Referensi
Rataan, Populasi	$\mu = \mu_0$	$\frac{(\bar{x} - \mu_0)}{s/\sqrt{n}}$	Z	#sampel besar (≥ 30)	statsmodels.stats.weightstats	https://www.statsmodels.org/dev/generated/statsmodels.stats.weightstats.ztest.html

Berikut merupakan hasil yang statistika yang didapatkan.

```
▶ 0.1s

1 statistik_man = (x_bar - 0.75)/(std_dev/(n**0.5))
2 print("Statistik:", statistik_man)
3 crit = stats.norm.ppf(0.99)
4 print(f"Critical region: ({crit},∞)")

Statistik: -9.780086778080644
Critical region: (2.3263478740408408,∞)
```

Dapat dilihat bahwa hasil statistik tidak jatuh pada critical region, maka dari itu Hipotesis null tidak ditolak sehingga data tidak cukup signifikan untuk menyimpulkan apapun.

Perhitungan Cara Modern Masalah 1

```
▶ 0.1s

1 stat, pval = statsmodels.stats.weightstats.ztest(df['Buta RLAC'], value=0.75,
alternative='larger')
2 print("Statistik:", stat)
3 print("P-value:", pval)

Statistik: -6.660562382406516
P-value: 0.999999999863609
```

Library python yang dipakai adalah statsmodel dan method yang digunakan adalah stats.weightstats.ztest. Dapat dilihat dari hasil diatas bahwa p value > alpha sehingga H0 tidak ditolak. Karena H0 tidak ditolak, maka data tidak cukup signifikan untuk menyimpulkan apapun.

Hal yang dapat dilakukan agar Hipotesis null ditolak adalah dengan menukar hipotesis. Hipotesis null sekarang adalah $H_0: \mu \geq 75\%$ dan Hipotesis Alternatif $H_1: \mu < 75\%$. Tingkat signifikansi yang digunakan masih sama yaitu 0.01.

Cara Manual (Setelah Hipotesis Masalah 1 Dibalik)

```
▶ 0.1s

1 statistik_man = (x_bar - 0.75)/(std_dev/(n**0.5))
2 print("Statistik:", statistik_man)
3 crit = stats.norm.ppf(0.99)
4 print(f"Critical region:  $(-\infty, \{crit\})$ ")

Statistik: -9.780086778080644
Critical region:  $(-\infty, 2.3263478740408408)$ 
```

Dapat dilihat, setelah hipotesis dibalik, maka hasil statistik sekarang berada di dalam critical region. Artinya, H_0 ditolak sehingga dapat disimpulkan bahwa dengan tingkat kepercayaan 99%, Mahasiswa CIT yang tidak melek Liberal Arts tidak lebih dari 75%.

Cara Modern (Setelah Hipotesis Masalah 1 Dibalik)

```
▶ 0.1s

1 stat, pval = statsmodels.stats.weightstats.ztest(df['Buta RLAC'], value=0.75,
2 alternative='smaller')
3 print("Statistik:", stat)
4 print("P-value:", pval)

Statistik: -6.660562382406516
P-value: 1.3639087576361523e-11
```

Dapat dilihat sekarang bahwa setelah hipotesis dibalik, p value lebih kecil dari alpha. Hal ini menyebabkan H_0 ditolak dan H_1 dapat dipertimbangkan. Maka dari itu, Dengan tingkat kepercayaan 99%, Mahasiswa CIT yang tidak melek Liberal Arts tidak lebih dari 75%.

Masalah 2:

Benarkah bahwa secara keseluruhan, besarnya persentase mahasiswa CIT yang tidak pernah sama sekali di sepanjang semester untuk meminjam atau membaca buku non-prodinya di perpustakaan lebih besar ketika dievaluasi pada semester ganjil ketimbang pada semester genap?

Misalkan:

X_1 adalah mahasiswa CIT yang membaca buku non-prodinya pada semester ganjil

X_2 adalah mahasiswa CIT yang membaca buku non-prodinya pada semester genap

Hipotesis null dan alternatif yang dipilih adalah sebagai berikut.

$$H_0: \mu_1 - \mu_2 \leq 0$$

$$H_1: \mu_1 - \mu_2 > 0$$

Kami memilih tingkat signifikansi = 0.01 untuk masalah ini dan statistik yang dipakai adalah selisih rata-rata dua populasi.

Pemrosesan Data

Sebelum masuk ke dalam perhitungan statistika terdapat pemrosesan data yang harus dilakukan.

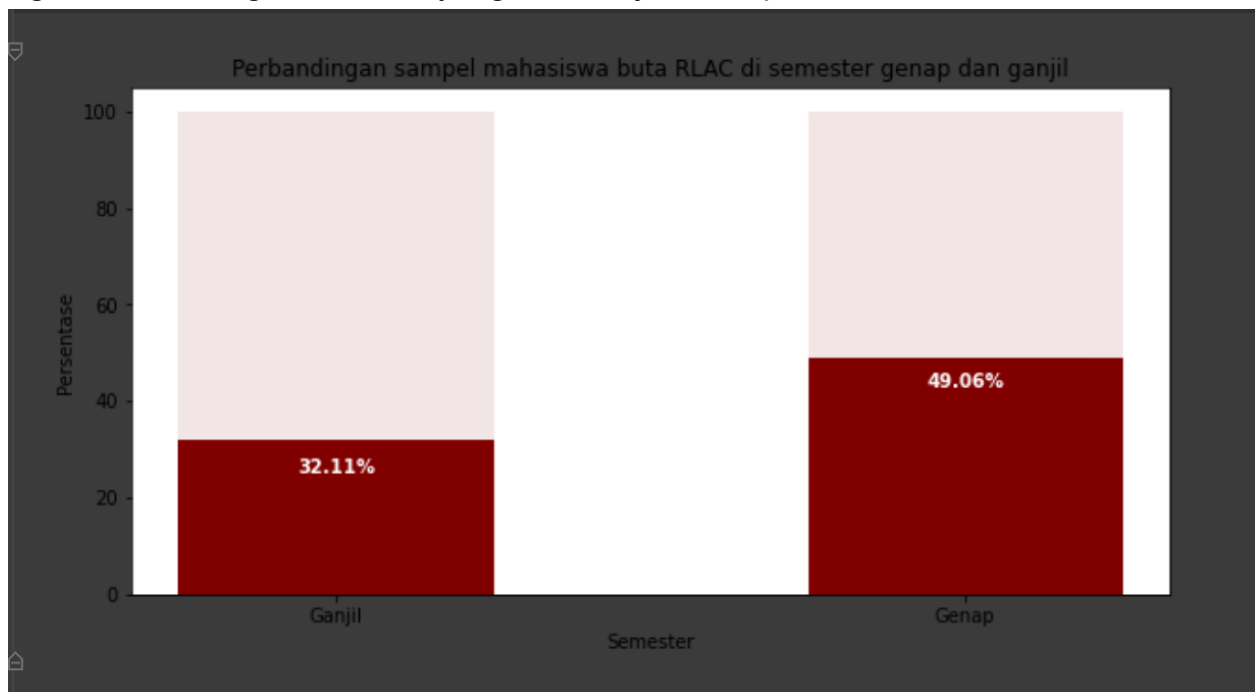

```
▶ 0.2s
1 tanggal_pinjam = df['Tanggal Pinjam']
2 semester_apa = []
3 for i in tanggal_pinjam:
4     tahun = i[-4:]
5     if tahun == '2021':
6         semester_apa.append('Ganjil')
7     else:
8         semester_apa.append('Genap')
9 semester_apa

'Ganjil',
'Ganjil',
'Ganjil',
'Ganjil',
'Ganjil',
'Ganjil',
'Ganjil',
'Ganjil',
'Ganjil',
'Ganjil',
```

Data sampel yang digunakan dalam masalah 2 adalah data sampel yang didapat dari akhir 2021 sampai awal 2022. Artinya, untuk memisahkan yang mana yang data semester ganjil dan yang mana data semester genap. Hal yang perlu dilakukan adalah melihat tahun nya, jika tahun nya adalah 2021, maka mahasiswa tersebut tergolong meminjam di semester ganjil, sebaliknya jika tahun nya adalah 2022, maka mahasiswa tersebut tergolong meminjam di semester genap.

```
▶ 0.5s
1 df2 = df
2 df2['Semester'] = semester_apa
3 df2 = df2.groupby("Nama Anggota").min()
4 df2
```

Setelah, dipisah berdasarkan semester, perlu dilakukan group by berdasarkan nama lagi untuk mencegah masalah yang sudah dijelaskan pada masalah 1.



Berikut merupakan visualisasi data dari masalah 2. Dapat dilihat bahwa terdapat 32.11 % mahasiswa di semester ganjil yang buta RLAC, sedangkan terdapat 49.06% mahasiswa di semester genap yang buta RLAC. Dari hal ini dapat diambil inferensi bahwa lebih banyak mahasiswa yang buta RLAC pada semester genap dibandingkan dengan semester ganjil. Hal ini juga dapat dibuktikan dengan hasil perhitungan statistika hipotesis dibawah ini.

Berikut merupakan statistika deskriptif dari masalah 2

```

▶ 0.4s

1 x_ganjil = np.mean(semester_ganjil)
2 x_genap = np.mean(semester_genap)
3 std_ganjil = semester_ganjil.std()
4 std_genap = semester_genap.std()
5 n1_df2 = len(semester_ganjil)
6 n2_df2 = len(semester_genap)
7 print(f"Rata - Rata Semester Ganjil = {x_ganjil}")
8 print(f"Rata - Rata Semester Genap = {x_genap}")
9 print(f"Standar Deviasi Semester Ganjil = {std_ganjil}")
10 print(f"Standar Deviasi Semester Genap = {std_genap}")

Rata - Rata Semester Ganjil = 0.3211009174311927
Rata - Rata Semester Genap = 0.49056603773584906
Standar Deviasi Semester Ganjil = 0.3211009174311927
Standar Deviasi Semester Genap = 0.49056603773584906

```

Cara Manual

Karena sampel yang dipilih ≥ 30 , maka data sampel termasuk data sampel yang besar. Maka dari itu, distribusi yang digunakan adalah distribusi normal (Z). Berikut merupakan perhitungan dengan cara manual. Rumus dan referensi library python yang digunakan adalah sebagai berikut.

Selisih Rataan 2 Populasi	$\mu_1 - \mu_2 = 0$	$\frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	Z	Keduanya #sampel besar (≥ 30)	statsmodels.stats.weightstats	https://www.statsmodels.org/dev/generated/statsmodels.stats.weightstats.ztest.html
---------------------------	---------------------	--	---	--------------------------------------	-------------------------------	---

```

▶ 0.4s

1 statistik_manual_df2 = (x_ganjil - x_genap) / (((std_ganjil**2)/n1_df2)+((std_genap**2)/n2_df2))**0.5
2 print("Hasil Statistik:", statistik_manual_df2)
3 crit = stats.norm.ppf(0.99)
4 print(f"Critical region:({crit} , ∞)")

Hasil Statistik: -2.051381175579928
Critical region:(2.3263478740408408 , ∞)

```

Dapat dilihat dari hasil statistik diatas bahwa hasil statistik berada diluar critical region. Maka dari itu, H_0 tidak ditolak sehingga data tidak cukup signifikan untuk menyimpulkan apapun.

Cara Modern

Berikut merupakan hasil perhitungan dengan cara modern.

```
1 tstat, pval = statsmodels.stats.weightstats.ztest(semester_ganjil, semester_genap, value=0, alternative='larger')
2 print(f'Hasil Statistik = {tstat}')
3 print(f'P-Value = {pval}')
```



```
Hasil Statistik = -2.1042314806479245
P-Value = 0.9823208701422311
```

Dapat dilihat dari hasil perhitungan diatas bahwa p value lebih besar dari alpha. Maka dari itu H_0 tidak ditolak sehingga data tidak cukup signifikan untuk menyimpulkan apapun.

Agar diperoleh hasil yang diinginkan yaitu H_0 ditolak, maka hal yang dapat dilakukan adalah untuk membalik hasil uji hipotesis. Sekarang uji hipotesis nya adalah sebagai berikut.

$$H_0: \mu_1 - \mu_2 \geq 0$$
$$H_1: \mu_1 - \mu_2 < 0$$

Cara Manual (Setelah Uji Hipotesis dibalik)

```
0.1s
1 statistik_manual_df2 = (x_ganjil - x_genap) / (((std_ganjil**2)/n1_df2)+((std_genap**2)/n2_df2))*0.5
2 print("Hasil Statistik:", statistik_manual_df2)
3 crit = stats.norm.ppf(0.99)
4 print(f"Critical region: (-∞ , {crit})")
```



```
Hasil Statistik: -2.051381175579928
Critical region: (-∞ , 2.3263478740408408)
```

Dapat dilihat dari hasil diatas bahwa hasil statistik berada di critical region, maka dari itu H_0 ditolak dan H_1 dapat dipertimbangkan. Maka dari itu, dapat ditarik kesimpulan bahwa dengan tingkat kepercayaan 99%, rata-rata mahasiswa yang buta RLAC pada semester ganjil tidak lebih besar dibandingkan dengan semester genap.

Cara Modern (Setelah Uji Hipotesis dibalik)

```
0.1s
1 tstat, pval = statsmodels.stats.weightstats.ztest(semester_ganjil, semester_genap, value=0, alternative='smaller')
2 print(f'Hasil Statistik = {tstat}')
3 print(f'P-Value = {pval}')
4 the_alpha_ = 0.01
5 if (pval < the_alpha_):
6     print('P Value < alpha (0.01)')
7 else:
8     print('P Value >= alpha (0.01)')

Hasil Statistik = -2.1042314806479245
P-Value = 0.017679129857768823
P Value >= alpha (0.01)
```

Dapat dilihat di perhitungan di atas bahwa dengan menggunakan method library statsmodel dari python yaitu statsmodels.stats.weightstats.ztest, p value masih \geq alpha. Hal ini berarti bahwa H_0 masih tidak ditolak. Kami masih bingung mengapa hasil yang diperoleh demikian, karena dengan menggunakan cara manual, setelah uji hipotesis dibalik, hasilnya menjadi H_0 ditolak. Maka dari itu, untuk mengatasi hal ini, kami menaikkan tingkat signifikansi menjadi 0.02

Berikut merupakan hasil perhitungan setelah hasil signifikansi dinaikkan.

```
0.1s
1 tstat, pval = statsmodels.stats.weightstats.ztest(semester_ganjil, semester_genap, value=0, alternative='smaller')
2 print(f'Hasil Statistik = {tstat}')
3 print(f'P-Value = {pval}')
4 the_alpha_ = 0.02
5 if pval <= the_alpha_:
6     print('H0 Ditolak')

Hasil Statistik = -2.1042314806479245
P-Value = 0.017679129857768823
H0 Ditolak
```

Dapat dilihat setelah menaikkan tingkat signifikansi, sekarang p value \leq alpha. Maka dari itu H_0 ditolak dan H_1 dapat dipertimbangkan. Maka dari itu, dapat ditarik kesimpulan bahwa dengan tingkat kepercayaan 98%, rata-rata mahasiswa yang buta RLAC pada semester ganjil tidak lebih besar dibandingkan dengan semester genap.

Masalah 3 :

Kode pre-prosessing data

```
df_th1 = pd.read_csv("angkatan_sulung.csv")
df_th1 = df_th1.groupby("Nama Anggota").min()
df_th1 = df_th1[["Buta RLAC"]]
df_th1.rename(columns={"Buta RLAC": "Buta RLAC tahun 1"}, inplace=True)
df_th1
```

Table Visualize Statistics

	Buta RLAC tahun 1	
Nama...		
Alfa ...	0	
Anata...	1	
Angel...	0	
Badia...	1	
Bamb...	0	
Brand...	1	
Chari...	0	
Christ...	0	

[27] ▶ 0.0s

```
df_th2 = df1[df1['angkatan'] == 19]
df_th2 = df_th2[["Buta RLAC"]]
df_th2.rename(columns={"Buta RLAC": "Buta RLAC tahun 2"}, inplace=True)
df_th2
```

Table Visualize Statistics

	Buta RLAC tahun 2	
Nama...		
Alfa ...	0	
Aquili...	0	
Badia...	0	
Brand...	0	

Data sampel untuk tahun pertama kami ambil dari data peminjaman buku angkatan pertama pada tahun pertamanya, dari awal semester ganjil (26-08-2019) sampai akhir semester genap (16-05-2020). Sedangkan untuk data sampel tahun kedua dari pertama kalinya perpustakaan buka di semester antara ditahun ke 3 angkatan pertama (30-09-2021) sampai tanggal 12-04-2022).

Kami membuat kolom "Buta RLAC" dan mengisinya secara manual dengan membandingkan judul buku dan prodi dari mahasiswa tersebut, "1" untuk buta RLAC dan "0" untuk melek RLAC. Jika mahasiswa pernah melakukan lebih dari satu peminjaman buku, kolom "Buta RLAC" akan ditandai dengan "0" jika setidaknya mahasiswa tersebut pernah sekali meminjam buku RLAC dan ditandai dengan "1" jika tidak pernah meminjam buku RLAC dalam 1 tahun akademik. Strategi tersebut kami lakukan dengan melakukan group by "nama" dan mengambil nilai paling kecil dari kolom "Buta RLAC".

Kami hanya mengambil kolom "nama" dan "Buta RLAC" untuk tahun pertama dan kedua. Kemudian dilakukan penggabungan inner join berdasarkan kolom nama agar hanya nama yang peminjaman yang terdaftar di kedua tahun yang diambil sebagai data pasangan. Dari 53 baris dari data tahun pertama dan 40 baris dari data tahun kedua, hasil penggabungan dengan inner join menghasilkan 34 baris data berpasangan.

```
[28] ▶ 0.4s
df_pasangan = pd.concat([df_th1, df_th2], axis=1, join="inner")
df_pasangan
```

	Buta RLAC tahun 1	Buta RLAC tahun 2
Nama...		
Alfa ...	0	0
Badia...	1	0
Brand...	1	0
Chari...	0	1
Christ...	0	0
Christ...	0	0
Christ...	0	0
Edgar...	1	0
Elisab...	0	0
Felici...	0	0

34 rows × 2 columns

Apakah benar bahwa mahasiswa CIT mengalami peningkatan literasi Liberal Arts ketika membandingkan pada pengalamannya di tahun pertama dan setelah tahun pertama?

$$H_1: \mu \geq 0$$

$H_2: \mu < 0$ (rata2 mahasiswa setelah tahun pertama lebih besar dari mahasiswa tahun pertama)



Dapat dilihat dari visualisasi diatas bahwa dari 33 sampel data, 20 orang tidak berubah, 4 orang memiliki peningkatan, dan 9 orang memiliki penurunan.

Misalkan

X_1 = mahasiswa tahun pertama X_2 = mahasiswa setelah tahun pertama

$$H_0: \mu d \leq 0$$

$H_1: \mu d > 0$ (mahasiswa mengalami peningkatan literasi liberal arts setelah tahun pertama)

Tingkat signifikansi = 0.1

Cara Modern

```
▶ 0.1s
1 statistic, pvalue = stats.ttest_rel(df_pasangan.iloc[:, 0], df_pasangan.iloc[:, 1], alternative="greater")
2 print(f'Hasil Statistik = {statistic}')
3 print(f'P Value = {pvalue}')
4 print(f'Tingkat Signifikansi = {0.1}')
5 if pvalue > 0.1:
6 |     print('H0 tidak ditolak')

Hasil Statistik = -1.4065622155408024
P Value = 0.9155474314301933
Tingkat Signifikansi = 0.1
H0 tidak ditolak
```

Dapat dilihat dari hasil diatas bahwa p value lebih besar dari tingkat signifikansi, maka dari itu H0 tidak ditolak dan hasil uji hipotesis tidak cukup untuk menyimpulkan apapun.

Membalik Uji Hipotesis

```
X1 = mahasiswa tahun pertama X2 = mahasiswa setelah tahun pertama

H0 :  $\mu \geq 0$ 
H1 :  $\mu < 0$  (mahasiswa mengalami penurunan literasi liberal arts setelah tahun pertama)
Tingkat signifikansi = 0.1
```

Cara Manual (Sebelum Dibalik)

```
▶ 0.1s
1 statistik_manual_df3 = (d_bar - 0) / (s/(n ** 0.5))
2 print("Hasil Statistik:", statistik_manual_df3)
3 hasil_ppf = stats.t.ppf(0.99, n - 1)
4 print(f'Critical Region: ({hasil_ppf}, ∞)')

Hasil Statistik: -1.4065622155408024
Critical Region: (2.4447941998077973, ∞)
```

Karena statistik tidak ada di dalam critical region maka kesimpulan Uji Hipotesis : H_0 tidak ditolak. Maka dari itu, data tidak cukup signifikan untuk menyimpulkan apapun.

Cara Manual (Setelah Uji Hipotesis Dibalik)

```
1 statistik_manual_df3 = (d_bar - 0) / (s/(n ** 0.5))
2 print("Hasil Statistik:", statistik_manual_df3)
3 hasil_ppf = stats.t.ppf(0.99, n - 1)
4 print(f'Critical Region: (-∞, {hasil_ppf})')

Hasil Statistik: -1.4065622155408024
Critical Region: (-∞, 2.4447941998077973)
```

Cara Modern (Setelah Uji Hipotesis Dibalik)

```
0.4s
1 statistic, pvalue = stats.ttest_rel(df_pasangan.iloc[:, 0], df_pasangan.iloc[:, 1], alternative="less")
2 print(f'Hasil Statistik = {statistic}')
3 print(f'P Value = {pvalue}')
4 print(f'Tingkat Signifikansi = {0.1}')
5 if pvalue <= 0.1:
6     print('H0 ditolak')

Hasil Statistik = -1.4065622155408024
P Value = 0.08445256856980668
Tingkat Signifikansi = 0.1
H0 ditolak
```

Dapat dilihat setelah hipotesis dibalik, maka sekarang p value lebih kecil dari tingkat signifikansi (0.1) maka dari itu H0 ditolak dan H1 dapat dipertimbangkan. Maka dari itu, dengan tingkat kepercayaan 90% mahasiswa CIT mengalami penurunan literasi liberal arts setelah tahun pertama.

Kesimpulan dan Saran

Kesimpulan

1. Persentase mahasiswa CIT yang tidak melek Liberal Arts kurang dari 75%.
2. Persentase mahasiswa buta RLAC pada semester ganjil lebih banyak daripada semester genap.
3. Mahasiswa CIT mengalami penurunan literasi liberal arts setelah tahun pertama.

Saran

- a. Data sampel yang diambil dapat lebih random.
- b. Jika dari hasil hipotesis tidak dapat ditarik kesimpulan apapun maka hal-hal berikut dapat dilakukan
 1. Tambah data sampel, semakin banyak data sampel maka akan semakin baik
 2. Tambahkan tingkat signifikansi, kurangi tingkat kepercayaan. Tingkat kepercayaan yang semakin rendah artinya range data yang semakin kecil, mungkin setelah tingkat kepercayaan diperkecil, kesimpulan uji hipotesis alternatif dapat dipertimbangkan.
 3. Menukar Hipotesis Null dengan Alternatif.
- ii. Gunakan Google untuk memastikan teori dan pemakaian syntax. Untuk kasus statistika hipotesis, tautan berikut dapat digunakan untuk mengecek hasil pekerjaan statistika.

<https://www.statssolver.com/hypothesis-testing.html>